

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 28, 2005, 14:12:12 ; Search time 167 Seconds
(without alignments)
551.191 Million cell updates/sec

Title: US-10-757-624-2-COPY
Perfect score: 1260
Sequence: 1 MSKBEELFTGVVPIVELDG.....VLKFTVAAGITGHMDELYK 238

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues
Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq.16Dec04:*
1: geneseqp19808:*
2: geneseqp19908:*
3: geneseqp20008:*
4: geneseqp20018:*
5: geneseqp20028:*
6: geneseqp20038:*
7: geneseqp20038s:*
8: geneseqp20048s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 1260 | 100.0 | 238 | 2 | AAW05304 |
| 2 | 1260 | 100.0 | 238 | 2 | AAW05312 |
| 3 | 1260 | 100.0 | 238 | 2 | AAW24232 |
| 4 | 1260 | 100.0 | 238 | 2 | AAW76110 |
| 5 | 1260 | 100.0 | 238 | 2 | AAW76106 |
| 6 | 1260 | 100.0 | 238 | 2 | AAW76113 |
| 7 | 1260 | 100.0 | 238 | 2 | AAW76105 |
| 8 | 1260 | 100.0 | 238 | 2 | AAW40479 |
| 9 | 1260 | 100.0 | 238 | 2 | AAW65079 |
| 10 | 1260 | 100.0 | 238 | 2 | AAW65078 |
| 11 | 1260 | 100.0 | 238 | 2 | AAW65080 |
| 12 | 1260 | 100.0 | 238 | 2 | AAW76371 |
| 13 | 1260 | 100.0 | 238 | 4 | AAW73552 |
| 14 | 1260 | 100.0 | 238 | 5 | AAE16046 |
| 15 | 1260 | 100.0 | 238 | 5 | AAE16042 |
| 16 | 1260 | 100.0 | 238 | 5 | AAE16043 |
| 17 | 1260 | 100.0 | 238 | 5 | AAE16045 |
| 18 | 1260 | 100.0 | 238 | 5 | AAE16073 |
| 19 | 1260 | 100.0 | 238 | 5 | AAE16044 |
| 20 | 1260 | 100.0 | 238 | 5 | AAE16047 |
| 21 | 1260 | 100.0 | 238 | 5 | AAE16041 |
| 22 | 1260 | 100.0 | 238 | 5 | AAE16087 |
| 23 | 1260 | 100.0 | 238 | 5 | AAE16038 |
| 24 | 1260 | 100.0 | 238 | 5 | ABG32365 |
| 25 | 1260 | 100.0 | 238 | 5 | ABG32367 |

| | | | | | | |
|----|------|-------|-----|---|----------|---------------------|
| 26 | 1260 | 100.0 | 238 | 5 | ABG32368 | Abg32368 Aequorea |
| 27 | 1260 | 100.0 | 238 | 6 | ABG76007 | Abg76007 Jellyfish |
| 28 | 1260 | 100.0 | 238 | 6 | ABG76008 | Abg76008 Jellyfish |
| 29 | 1260 | 100.0 | 238 | 6 | ABG75980 | Abg75980 Jellyfish |
| 30 | 1260 | 100.0 | 238 | 6 | ABG34939 | Aeg34939 Aequorea |
| 31 | 1260 | 100.0 | 238 | 6 | ABR44423 | Abri44423 Wild-type |
| 32 | 1260 | 100.0 | 238 | 6 | ADA25221 | Ada25221 Green flu |
| 33 | 1260 | 100.0 | 238 | 6 | ADA25216 | Ada25216 Green flu |
| 34 | 1260 | 100.0 | 238 | 6 | ADA25212 | Ada25212 Green flu |
| 35 | 1260 | 100.0 | 238 | 6 | ADA25215 | Ada25215 Green flu |
| 36 | 1260 | 100.0 | 238 | 6 | ADA25196 | Ada25196 Green flu |
| 37 | 1260 | 100.0 | 238 | 6 | ADA25214 | Ada25214 Green flu |
| 38 | 1260 | 100.0 | 238 | 6 | ADA25194 | Ada25194 Aequorea |
| 39 | 1260 | 100.0 | 238 | 6 | ADA25218 | Ada25218 Green flu |
| 40 | 1260 | 100.0 | 238 | 6 | ADA25195 | Ada25195 Green flu |
| 41 | 1260 | 100.0 | 238 | 7 | ADF70380 | Adf70380 Aequorea |
| 42 | 1260 | 100.0 | 238 | 7 | ADM78505 | Adm78505 Wild-type |
| 43 | 1260 | 100.0 | 238 | 7 | ADM78577 | Adm78577 Mutant Ae |
| 44 | 1260 | 100.0 | 238 | 8 | ADQ59552 | Adq59552 Aequorea |
| 45 | 1260 | 100.0 | 501 | 2 | AAW31878 | Aaw31878 GFP varia |

ALIGNMENTS

RESULT 1
ID AAW05304 standard; protein; 238 AA.
XX AAW05304;
AC
XX
DT 02-APR-1997 (first entry)
XX
DE Green fluorescent protein.
XX
KM Green fluorescent protein; GFP; jellyfish; Aequorea victoria; sea pansy;
KM Renilla reniformis; differential gene expression; protein localisation;
KM gene expression tracking; fluorescence.
XX
OS Aequorea victoria.
XX
PN WO9623810-A1.
XX
PD 08-AUG-1996.
XX
PF 13-NOV-1995; 95WO-US014692.
XX
PR 10-NOV-1994; 94US-00337915.
XX
PA (REGC) UNITV CALIFORNIA.
XX
PI Tsien RY, Heim R;
XX
DR WPI; 1996-371370/37.
DR N-PSDB; AAT39694.
XX
PT New modified Aequorea green fluorescent polypeptide(s) - having amino
PT acid changes to provide prods. which exhibit different excitation and
PT emission spectra.
PS
PS Disclosure; Page 20-21; 39pp; English.
XX
CC This sequence represents the wild type Aequorea victoria (jellyfish)
CC green fluorescent protein (GFP). The fluorescence of this protein is
CC generated by cyclisation and oxidation of the Ser-Tyr-Gly sequence at
CC positions 65-67. Aequorea GFP has two absorbance peaks, as opposed to the
CC one absorbance peak seen in the related GFP from the sea pansy (Renilla
CC reniformis). This protein is modified, to provide the GFPs of the
CC invention. The modifications were created by subjecting the cDNA encoding
CC this sequence to site directed mutagenesis using mutagenic PCR primers,
CC or hydroxylamine treatment. The GFPs of the invention are modified to
CC lead to the formation of products with markedly different excitation and
CC emission spectra. Visibly distinct colours, and increased intensities of

CC emission make these products useful in a wide variety of contexts, such
CC as tracking of differential gene expression and protein localisation. The
CC mutations can also be created to modify the encoded GFP so that it only
CC possesses one absorbance peak

XX Sequence 238 AA;

Query Match 100.0%; Score 1260; DB 2; Length 238;
Best Local Similarity 98.3%; Pred. No. 1.5e-119;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVPIVLELDGNGHKFSVSGEGDATYGLTKLTKFTCTTGKLPVPMPTL 60
Db 1 MSKGEELFTGVPIVLELDGNGHKFSVSGEGDATYGLTKLTKFTCTTGKLPVPMPTL 60

QY VTTXXYGVOCFSRYPDMKRRHDFPKSAMPEGVVOERTIFFKDDGNYKTRAAYVFBGDTLV 120
Db 61 VTTXXYGVOCFSRYPDMKRRHDFPKSAMPEGVVOERTIFFKDDGNYKTRAAYVFBGDTLV 120

QY 121 NRLEKIDPFKEDGNILGHKLEYNNSHNYIMADKQKGIKVNFKIRHNIEDGXVOLAD 180
Db 121 NRLEKIDPFKEDGNILGHKLEYNNSHNYIMADKQKGIKVNFKIRHNIEDGXVOLAD 180

QY 181 HYQONTPIGDGPVLLPDNNHILSTQSALSQDPNEKRDHVLKFEVTAAGITTHGMDELYK 238
Db 181 HYQONTPIGDGPVLLPDNNHILSTQSALSQDPNEKRDHVLKFEVTAAGITTHGMDELYK 238

RESULT 2

AAW05312
ID AAW05312 standard; protein; 238 AA.

AC AAW05312;

DT 02-Apr-1997 (first entry)

DE Green fluorescent protein mutant S65X.

KW Green fluorescent protein; GFP; jellyfish; Aequorea victoria; sea pansy;
KM Renilla reniformis; differential gene expression; protein localisation;
KW gene expression tracking; fluorescence.

OS Synthetic.

Key Location/Qualifiers
FT Misc-difference 65 /label= Ala, Cys, Thr, Leu, Val, Ile
FT /note= "S65X"

PN MO9623810-A1.

PD 08-AUG-1996.

PF 13-NOV-1995; 95WO-US014692.

PR 10-NOV-1994; 94US-00337915.

XX (REGC) UNIV CALIFORNIA.

XX Teien RY, Heim R;

DR WPI; 1996-371370/37.

XX New modified Aequorea green fluorescent polypeptide(s) - having amino
PT acid changes to provide prode. which exhibit different excitation and
XX emission spectra.

PS Claim 14; Page 7; 39pp; English.

XX AAW05305-W05312 represent Aequorea victoria (jellyfish) green fluorescent
CC protein (GFP) of the invention (see AAW05304 for the wild type protein).
CC The fluorescence of this protein is generated by cyclisation and
CC oxidation of the Ser-Tyr-Gly sequence at positions 65-67. Aequorea GFP

CC has two absorbance peaks, as opposed to the one absorbance peak seen in
CC the related GFP from the sea pansy (Renilla reniformis). The
CC modifications present in these sequences were created by subjecting the
CC cDNA encoding this sequence to site directed mutagenesis using mutagenic
CC PCR primers, or hydroxylamine treatment. These GFPs of the invention are
CC modified to lead to the formation of products with markedly different
CC excitation and emission spectra. Visibly distinct colours, and increased
CC intensities of emission make these products useful in a wide variety of
CC contexts, such as tracking of differential gene expression and protein
CC localisation. The mutations can also be created to modify the encoded GFP
CC so that it only possesses one absorbance peak

XX Sequence 238 AA;

Query Match 100.0%; Score 1260; DB 2; Length 238;
Best Local Similarity 98.7%; Pred. No. 1.5e-119;
Matches 235; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVPIVLELDGNGHKFSVSGEGDATYGLTKLTKFTCTTGKLPVPMPTL 60
Db 1 MSKGEELFTGVPIVLELDGNGHKFSVSGEGDATYGLTKLTKFTCTTGKLPVPMPTL 60

QY 61 VTTXXYGVOCFSRYPDMKRRHDFPKSAMPEGVVOERTIFFKDDGNYKTRAAYVFBGDTLV 120
Db 61 VTTXXYGVOCFSRYPDMKRRHDFPKSAMPEGVVOERTIFFKDDGNYKTRAAYVFBGDTLV 120

QY 121 NRLEKIDPFKEDGNILGHKLEYNNSHNYIMADKQKGIKVNFKIRHNIEDGXVOLAD 180
Db 121 NRLEKIDPFKEDGNILGHKLEYNNSHNYIMADKQKGIKVNFKIRHNIEDGXVOLAD 180

QY 181 HYQONTPIGDGPVLLPDNNHILSTQSALSQDPNEKRDHVLKFEVTAAGITTHGMDELYK 238
Db 181 HYQONTPIGDGPVLLPDNNHILSTQSALSQDPNEKRDHVLKFEVTAAGITTHGMDELYK 238

RESULT 3

AAW24232
ID AAW24232 standard; protein; 238 AA.

AC AAW24232;

DT 02-FEB-1998 (first entry)

DE Aequorea victoria green fluorescent protein (GFP).

KW North West Pacific jellyfish; green fluorescent protein; GFP; S65C;
KW fluorescent resonance energy transfer; FRET; enzymatic assay;
KW enzyme inhibitor; enzyme promoter; Tandem fluorescent protein construct;
KW blue fluorescent protein; protease cleavage site.

XX Aequorea victoria.

PN MO9728261-A1.

PD 07-AUG-1997.

PF 31-JAN-1997; 97WO-US001457.

PR 31-JAN-1996; 96US-00594575.

XX (REGC) UNIV CALIFORNIA.

XX (AURO-) AURORA BIOSCIENCES CORP.

XX Teien RY, Heim R, Cubitt A;

DR WPI; 1997-402615/37.

DR N-PSDB; AAT77843.

XX Tandem fluorescent protein constructs - have donor and acceptor moieties
PT exhibiting fluorescent linked via cleavable peptide linker, useful in
XX enzymatic assays.

PS Disclosure; Fig 1; 88pp; English.

XX This sequence represents Aequorea victoria (North West Pacific jellyfish)
 CC green fluorescent protein (GFP). The GFP protein is used in a novel
 CC tandem fluorescent protein construct. The construct comprises a donor
 CC (e.g. GFP protein) and an acceptor (e.g. S65C, a variant of GFP or blue
 CC fluorescent protein) fluorescent protein moiety, and a linker coupling
 CC them. Preferably, the donor is positioned at the N-terminus of the
 CC polypeptide relative to the acceptor. The linker moiety is a peptide 5-50
 CC amino acids in length containing a protease cleavage site. The donor and
 CC acceptor moieties exhibit fluorescent resonance energy transfer (FRET)
 CC when the donor moiety is excited. The constructs are used in enzymatic
 CC assays and can be used to isolate new enzymes or enzyme inhibitors or
 CC promoters. The specific activity of enzyme (in vivo and in vitro) and
 CC compounds altering enzyme activity can be obtained. FRET and hence
 CC activity of specific compounds is measured from the acceptor or donor
 CC moiety or maybe obtained using a ratio between the two
 CC
 SQ Sequence 238 AA;

Query Match 100.0%; Score 1260; DB 2; Length 238;
 Best Local Similarity 98.3%; Pred. No. 1.5e-119;
 Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGELFTGVVPIVELDGDVNGHKFSVSGEGDATYGLTLKFTCTTGKLPVWPPTL 60
 DB 1 MSKGELFTGVVPIVELDGDVNGHKFSVSGEGDATYGLTLKFTCTTGKLPVWPPTL 60
 QY 61 VTTXXYGVQCSRRYPDHMKRHDFFKSAPEGVQERTIFFKDDGNYKTRAEVKFGDITLV 120
 DB 61 VTTFSYGVQCSRRYPDHMKRHDFFKSAPEGVQERTIFFKDDGNYKTRAEVKFGDITLV 120
 QY 121 NRIELKGIQDKEDNGIILHKLKLEVYNNSHNVYIMADKQNGIKVNPKRININIDGVVOLAD 180
 DB 121 NRIELKGIQDKEDNGIILHKLKLEVYNNSHNVYIMADKQNGIKVNPKRININIDGVVOLAD 180
 QY 181 HYQONTPIGDGPVLLPDNHYLSTOSALSKDPNEKRDHVLKLVTAAGITTHGMDELYK 238
 DB 181 HYQONTPIGDGPVLLPDNHYLSTOSALSKDPNEKRDHVLKLVTAAGITTHGMDELYK 238

RESULT 4
 AAM76110
 ID AAM76110 standard; protein; 238 AA.
 XX
 AC AAM76110;

DT 18-NOV-1998 (first entry)
 XX
 DE A. victoria green fluorescent protein mutant S65A.
 XX
 KM Green fluorescent protein; GFP; mutant; jellyfish; excitation;
 XX chromophore.
 XX
 OS Aequorea victoria.
 XX
 OS Synthetic.

FT Key Location/Qualifiers
 FT Misc-difference 65
 FT /label= S65A
 FT /note= "Wild-type Ser is replaced by Ala"

XX US5804387-A.
 XX
 PD 08-SEP-1998.
 XX
 PF 31-JAN-1997; 97US-00791332.
 XX
 PR 01-FEB-1996; 96US-0010960P.
 XX
 PA (STRD) UNIV IELAND STANFORD JUNIOR.
 XX
 PI Cormack BP, Valdivia RH, Falkow S;
 XX

DR WPI, 1998-505643/43.
 XX
 XX DNA encoding mutant green fluorescent pigment proteins - with greater
 PT fluorescence intensity than wild-type proteins, useful for studying gene
 PT expression and protein localisation.
 XX
 XX
 PS Claim 6; Page; 15pp; English.

XX This sequence represents a mutant green fluorescent protein (GFP) from
 CC Aequorea victoria in which a Ser residue at position 65 in the wild type
 CC protein is replaced by an Ala. These mutations occur in the chromophore
 CC region. GFP mutants are used in a method to discover GFP's mutated in the
 CC chromophore region that fluoresce more brightly than wild-type GFP upon
 CC excitation at 488 nm. These mutants can be used in a method for analyzing
 CC a cell containing the GFP. The GFP can also be fused to a protein and
 CC used to identify the intracellular localisation of a protein of interest.
 CC A regulatory element could be operatively connected to a coding portion
 CC encoding a mutant GFP and exposed to an environmental stimulus. The
 CC fluorescence signal from the GFP then measures the effect of
 CC the stimulus on the regulatory element. NOTE: This sequence does not
 CC appear in the specification but has been constructed from the wild-type
 CC GFP protein represented in AAM76106
 CC
 SQ Sequence 238 AA;

Query Match 100.0%; Score 1260; DB 2; Length 238;
 Best Local Similarity 98.3%; Pred. No. 1.5e-119;
 Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGELFTGVVPIVELDGDVNGHKFSVSGEGDATYGLTLKFTCTTGKLPVWPPTL 60
 DB 1 MSKGELFTGVVPIVELDGDVNGHKFSVSGEGDATYGLTLKFTCTTGKLPVWPPTL 60
 QY 61 VTTXXYGVQCSRRYPDHMKRHDFFKSAPEGVQERTIFFKDDGNYKTRAEVKFGDITLV 120
 DB 61 VTTFSYGVQCSRRYPDHMKRHDFFKSAPEGVQERTIFFKDDGNYKTRAEVKFGDITLV 120
 QY 121 NRIELKGIQDKEDNGIILHKLKLEVYNNSHNVYIMADKQNGIKVNPKRININIDGVVOLAD 180
 DB 121 NRIELKGIQDKEDNGIILHKLKLEVYNNSHNVYIMADKQNGIKVNPKRININIDGVVOLAD 180
 QY 181 HYQONTPIGDGPVLLPDNHYLSTOSALSKDPNEKRDHVLKLVTAAGITTHGMDELYK 238
 DB 181 HYQONTPIGDGPVLLPDNHYLSTOSALSKDPNEKRDHVLKLVTAAGITTHGMDELYK 238

RESULT 5
 AAM76106
 ID AAM76106 standard; protein; 238 AA.
 XX
 AC AAM76106;

DT 18-NOV-1998 (first entry)
 XX
 DE A. victoria green fluorescent protein mutant F64L/S65T.
 XX
 KM Green fluorescent protein; GFP; mutant; jellyfish; excitation;
 XX chromophore.
 XX
 OS Aequorea victoria.
 XX
 OS Synthetic.

FT Key Location/Qualifiers
 FT Misc-difference 64
 FT /label= F64L
 FT /note= "Wild-type Phe is replaced with Leu"
 FT Misc-difference 65
 FT /label= S65T
 FT /note= "Wild-type Ser is replaced by Thr"

XX US5804387-A.
 XX
 PD 08-SEP-1998.
 XX

XX 31-JAN-1997; 97US-00791332.
XX PF Aequorea victoria in which a Phe residue at position 64 in the wild type
XX PR 01-FEB-1996; 96US-0010960P.
XX (STRD) UNIV LELAND STANFORD JUNIOR.
XX PI Cormack BP, Valdivia RH, Falkow S;
XX WPI, 1998-505643/43.
XX
XX DNA encoding mutant green fluorescent pigment proteins - with greater
XX fluorescence intensity than wild-type proteins, useful for studying gene
XX expression and protein localisation.
XX
XX PS Disclosure; Page; 15pp; English.
XX
XX This sequence represents a mutant green fluorescent protein (GFP) from
XX Aequorea victoria in which a Phe residue at position 64 in the wild type
XX protein is replaced by a Leu and a Ser at position 65 is replaced by a
XX Thr. These mutations occur in the chromophore region. GFP mutants are
XX used in a method to discover GFP's mutated in the chromophore region that
XX fluoresce more brightly than wild-type GFP upon excitation at 488 nm.
XX These mutants can be used in a method for analyzing a cell containing the
XX GFP. The GFP can also be fused to a protein of interest. A regulatory element
XX could be operatively connected to a coding portion encoding a mutant GFP
XX the from the cell then measures the effect of the stimulus on the
XX regulatory element. NOTE: This sequence does not appear in the
XX CC specification but has been constructed from the wild-type GFP protein
XX represented in AAW76106

XX SQ Sequence 238 AA;

Query Match 100.0%; Score 1260; DB 2; Length 238;
Best Local Similarity 98.3%; Pred. No. 1.5e-119;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEEFTGVVPIVLELDGVNGHKFSVSGEGDATTGKLTFRCTTGKLPVWPPTL 60
DB 1 MSKGEEFTGVVPIVLELDGVNGHKFSVSGEGDATTGKLTFRCTTGKLPVWPPTL 60
QY VTTXXYGVCCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAIVKFEEDTLV 120
DB 61 VTTLTGVCCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAIVKFEEDTLV 120
QY 121 NRTELKIDFKEDGNILGHKLEYNNSHNVYIMADKXNGIKVNFKIRNIEDGXVQLAD 180
DB 121 NRTELKIDFKEDGNILGHKLEYNNSHNVYIMADKXNGIKVNFKIRNIEDGXVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLFVTAAGITHGMDLEYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLFVTAAGITHGMDLEYK 238

RESULT 6

AAW76113
ID AAW76113 standard; protein; 238 AA.

XX AAW76113;

XX 18-NOV-1998 (first entry)

DE A. victoria green fluorescent protein mutant S65G.

KW Green fluorescent protein; GFP; mutant; jellyfish; excitation;
chromophore.

XX Aequorea victoria.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 65
FT /label= S65G
FT /note= "Wild-type Ser is replaced by Gly"

XX US5804387-A.

XX 08-SEP-1998.

XX 31-JAN-1997; 97US-00791332.

XX 01-FEB-1996; 96US-0010960P.

XX (STRD) UNIV LELAND STANFORD JUNIOR.

XX PI Cormack BP, Valdivia RH, Falkow S;

XX WPI, 1998-505643/43.

XX DNA encoding mutant green fluorescent pigment proteins - with greater
XX fluorescence intensity than wild-type proteins, useful for studying gene
XX expression and protein localisation.

XX Claim 9; Page; 15pp; English.

XX This sequence represents a mutant green fluorescent protein (GFP) from
XX Aequorea victoria in which a Gly. These mutations occur in the chromophore
XX region. GFP mutants are used in a method to discover GFP's mutated in the
XX chromophore region that fluoresce more brightly than wild-type GFP upon
XX excitation at 488 nm. These mutants can be used in a method for analyzing
XX a cell containing the GFP. The GFP can also be fused to a protein and
XX used to identify the intracellular localisation of a protein of interest.
XX A regulatory element could be operatively connected to a coding portion
XX encoding a mutant GFP and exposed to an environmental stimulus. The
XX fluorescence signal from the cell then measures the effect of
XX the stimulus on the regulatory element. NOTE: This sequence does not
XX appear in the specification but has been constructed from the wild-type
XX GFP protein represented in AAW76106

XX SQ Sequence 238 AA;

Query Match 100.0%; Score 1260; DB 2; Length 238;
Best Local Similarity 98.3%; Pred. No. 1.5e-119;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEEFTGVVPIVLELDGVNGHKFSVSGEGDATTGKLTFRCTTGKLPVWPPTL 60
DB 1 MSKGEEFTGVVPIVLELDGVNGHKFSVSGEGDATTGKLTFRCTTGKLPVWPPTL 60
QY VTTXXYGVCCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAIVKFEEDTLV 120
DB 61 VTTGTGVCCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAIVKFEEDTLV 120
QY 121 NRTELKIDFKEDGNILGHKLEYNNSHNVYIMADKXNGIKVNFKIRNIEDGXVQLAD 180
DB 121 NRTELKIDFKEDGNILGHKLEYNNSHNVYIMADKXNGIKVNFKIRNIEDGXVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLFVTAAGITHGMDLEYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLFVTAAGITHGMDLEYK 238

RESULT 7

AAW76105
ID AAW76105 standard; protein; 238 AA.

XX AAW76105;

XX 18-NOV-1998 (first entry)

DE A. victoria green fluorescent protein.

KW Green fluorescent protein; GFP; mutant; jellyfish; excitation;

KW chromosome.
 OS Aequorea victoria.
 XX US804387-A.
 XX 08-SEP-1998.
 XX 31-JAN-1997; 97US-00791332.
 XX 01-FEB-1996; 96US-0010960P.
 XX (STRD) UNIV LELAND STANFORD JUNIOR.
 XX Cormack BP, Valdivia RH, Falkow S;
 XX WPI; 1998-505643/43.
 XX
 XX DNA encoding mutant green fluorescent pigment proteins - with greater
 PT fluorescence intensity than wild-type proteins, useful for studying gene
 PT expression and protein localization.
 XX
 XX Disclosure; Page; 15pp; English.
 XX
 XX This sequence represents the green fluorescent protein (GFP) from
 CC Aequorea victoria. This sequence is used in a method to discover GFP's
 CC mutated in the chromophore region that fluoresce more brightly than wild-
 CC type GFP upon excitation at 488 nm. These mutants can be used in a method
 CC for analyzing a cell containing the GFP. The GFP can also be fused to a
 CC protein and used to identify the intracellular localization of a protein
 CC of interest. A regulatory element could be operatively connected to a
 CC coding portion encoding a mutant GFP and exposed to an environmental
 CC stimulus. The fluorescence signal from the GFP from the cell then measures
 CC the effect of the stimulus on the regulatory element. Note: This sequence
 CC does not appear in the specification but has been retrieved from the EMBL
 CC record contained in W05304
 XX
 XX Sequence 238 AA;
 SQ
 Query Match 100.0%; Score 1260; DB 2; Length 238;
 Best Local Similarity 98.3%; Pred. No. 1.5e-119;
 Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSKGELFTGVVPIVELDGVNGHKFSVSGEGDATYGLTKLFTCTTGLPVPWPTL 60
 DB 1 MSKGELFTGVVPIVELDGVNGHKFSVSGEGDATYGLTKLFTCTTGLPVPWPTL 60
 QY VTTXXYGVQCFSRYPDHMKRHDFFKSAPEGVQERTIFFKODGNYKTRAEVKFECDTLV 120
 DB 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAPEGVQERTIFFKODGNYKTRAEVKFECDTLV 120
 QY 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
 DB 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
 QY 181 HYQONTPIGDGVVLLPDNHYLSTQSALSKDPNEKRDHVLLEFYTAAGITTHGMDLYK 238
 DB 181 HYQONTPIGDGVVLLPDNHYLSTQSALSKDPNEKRDHVLLEFYTAAGITTHGMDLYK 238
 RESULT 8
 AAM40479
 ID AAM40479 standard; protein; 238 AA.
 XX
 XX AAM40479;
 AC
 XX 15-SEP-1998 (first entry)
 DT
 XX
 XX A. victoria green fluorescent protein.
 DE
 XX Green fluorescent protein; protein kinase; phosphate donor;
 KW phosphorylation; drug screening; receptor-ligand binding; signalling;
 KW protein-protein interaction; kinase activation.

XX Aequorea victoria.
 OS
 XX Key Location/Qualifiers
 FH 1.238
 FT Protein /note="partial sequence"
 XX
 XX W09802571-A1.
 XX 22-JAN-1998.
 XX 16-JUL-1997; 97WO-US012410.
 XX 16-JUL-1996; 96US-00679865.
 XX 16-JUL-1996; 96US-00680876.
 XX 16-JUL-1996; 96US-00680877.
 XX (REGC) UNIV CALIFORNIA.
 XX
 XX Tsien RY, Cudde AB;
 XX WPI; 1998-110616/10.
 XX N-PSDB; AAV11125.
 XX
 XX Assays for protein kinase and modulators - using a fluorescent protein
 PT substrate which exhibits a different fluorescent property in the
 PT phosphorylated and un-phosphorylated state.
 XX
 XX Disclosure; Fig 3; 65pp; English.
 XX
 XX This sequence represents a green fluorescent protein from Aequorea
 CC victoria which is used in a method to determine if a sample contain
 CC protein kinase (PK) activity. The method involves contacting the sample
 CC with a phosphate donor (PD) and a fluorescent protein (FP) substrate for
 CC a PK, the protein substrate comprising a FP moiety and a phosphorylation
 CC site for a PK, where the protein substrate exhibits a different
 CC fluorescent property in the phosphorylated state than in the un-
 CC phosphorylated state. The protein substrate is then excited and the
 CC amount of a fluorescent property that differs in the un-phosphorylated
 CC state and phosphorylated state is measured, whereby an amount that is
 CC consistent with the presence of the protein substrate in its
 CC phosphorylated state indicates the presence of PK activity. The method
 CC and products can be used in drug screening. They can be used for
 CC screening for compounds which affect cellular events, including receptor-
 CC ligand binding, protein-protein interactions or kinase activation, which
 CC signal to the target kinase
 XX
 XX Sequence 238 AA;
 SQ
 Query Match 100.0%; Score 1260; DB 2; Length 238;
 Best Local Similarity 98.3%; Pred. No. 1.5e-119;
 Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSKGELFTGVVPIVELDGVNGHKFSVSGEGDATYGLTKLFTCTTGLPVPWPTL 60
 DB 1 MSKGELFTGVVPIVELDGVNGHKFSVSGEGDATYGLTKLFTCTTGLPVPWPTL 60
 QY VTTXXYGVQCFSRYPDHMKRHDFFKSAPEGVQERTIFFKODGNYKTRAEVKFECDTLV 120
 DB 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAPEGVQERTIFFKODGNYKTRAEVKFECDTLV 120
 QY 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
 DB 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
 QY 181 HYQONTPIGDGVVLLPDNHYLSTQSALSKDPNEKRDHVLLEFYTAAGITTHGMDLYK 238
 DB 181 HYQONTPIGDGVVLLPDNHYLSTQSALSKDPNEKRDHVLLEFYTAAGITTHGMDLYK 238
 RESULT 9
 AAM65079
 ID AAM65079 standard; protein; 238 AA.

```

XX AC AAM65078;
XX DT 15-SEP-1998 (first entry)
XX DE A. victoria green fluorescent protein variant S65A.
XX KM Green fluorescent protein; protein kinase; phosphate donor;
XX KW phosphorylation; drug screening; receptor-ligand binding; signalling;
XX protein-protein interaction; kinase activation.
XX OS Aequorea victoria.
XX SS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 65 /label= S65A
XX FT /note= "wild type Ser is replaced with Ala"
XX PN MO9802571-Al.
XX PD 22-JAN-1998.
XX PF 16-JUL-1997; 97WO-US012410.
XX PR 16-JUL-1996; 96US-00679865.
XX PR 16-JUL-1996; 96US-00680876.
XX PR 16-JUL-1996; 96US-00680877.
XX PA (REGC ) UNIV CALIFORNIA.
XX PI Tsien RY, Cubitt AB;
XX DR WPI; 1998-110616/10.
XX PT Assays for protein kinase and modulators - using a fluorescent protein
XX PT substrate which exhibits a different fluorescent property in the
XX PT phosphorylated and un-phosphorylated state.
XX PS Claim 27; Page; 65pp; English.
XX CC This sequence represents a green fluorescent protein variant from
XX CC Aequorea victoria in which a wild-type Ser residue at position 65 is
XX CC replaced with a Ala. This variant is used in a method to determine if a
XX CC sample contains protein kinase (PK) activity. The method involves
XX CC contacting the sample with a phosphate donor (PD) and a fluorescent
XX CC protein (FP) substrate for a PK, the protein substrate comprising a FP
XX CC moiety and a phosphorylation site for a PK, where the protein substrate
XX CC exhibits a different fluorescent property in the phosphorylated state
XX CC than in the un-phosphorylated state. The protein substrate is then
XX CC excited and the amount of a fluorescent property that differs in the un-
XX CC phosphorylated state and phosphorylated state is measured, whereby an
XX CC amount that is consistent with the presence of the protein substrate in
XX CC its phosphorylated state indicates the presence of PK activity. The
XX CC method and products can be used in drug screening. They can be used for
XX CC screening for compounds which affect cellular events, including receptor-
XX CC ligand binding, protein-protein interactions or kinase activation, which
XX CC signal to the target kinase. NOTE: This sequence does not appear in the
XX CC specification and has been constructed from the wild-type sequence
XX CC represented in AAM40479
XX SQ Sequence 238 AA;
XX
XX Query Match 100.0%; Score 1360; DB 2; Length 238;
XX Best Local Similarity 98.3%; Pred. No. 1.5e-119;
XX Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 MSKGERLFTGVPIIVIELDGVNNGHFRSVSGSEGGATYCKTLTKRTCTTGKLPVWPMTL 60
XX Db 1 MSKGERLFTGVPIIVIELDGVNNGHFRSVSGSEGGATYCKTLTKRTCTTGKLPVWPMTL 60
XX 1 MSKGERLFTGVPIIVIELDGVNNGHFRSVSGSEGGATYCKTLTKRTCTTGKLPVWPMTL 60
XX 61 VTTXXYGVOCFSRYPDMKRGHDFPKSAMPEGYVOERTIFFKXDGNAVKTAEVFGSDTLV 120
XX ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 61 VTTFAYGVOCFSRYPDMKRGHDFPKSAMPEGYVOERTIFFKXDGNAVKTAEVFGSDTLV 120
XX
XX Qy 121 NRLEKGIIDFKEDGNILGHLENTYNSHNVYIMADKQKNGIKRVNFKIRNIEDGXVOLAD 180
XX ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX Db 121 NRLEKGIIDFKEDGNILGHLENTYNSHNVYIMADKQKNGIKRVNFKIRNIEDGXVOLAD 180
XX
XX Qy 181 HYQONTPIGGPVLTPDNNHYLSTQSALSKDPNEKRDHMLTKFVTAAAGITHGDELTX 238
XX ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX Db 181 HYQONTPIGGPVLTPDNNHYLSTQSALSKDPNEKRDHMLTKFVTAAAGITHGDELTX 238
XX
XX RESULT 10
XX ID AAM65078 standard; protein; 238 AA.
XX AC AAM65078;
XX DT 15-SEP-1998 (first entry)
XX DE A. victoria green fluorescent protein variant S65T.
XX KM Green fluorescent protein; protein kinase; phosphate donor;
XX KW phosphorylation; drug screening; receptor-ligand binding; signalling;
XX protein-protein interaction; kinase activation.
XX OS Aequorea victoria.
XX SS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 65 /label= S65T
XX FT /note= "wild type Ser is replaced with Thr"
XX PN MO9802571-Al.
XX PD 22-JAN-1998.
XX PF 16-JUL-1997; 97WO-US012410.
XX PR 16-JUL-1996; 96US-00679865.
XX PR 16-JUL-1996; 96US-00680876.
XX PR 16-JUL-1996; 96US-00680877.
XX PA (REGC ) UNIV CALIFORNIA.
XX PI Tsien RY, Cubitt AB;
XX DR WPI; 1998-110616/10.
XX PT Assays for protein kinase and modulators - using a fluorescent protein
XX PT substrate which exhibits a different fluorescent property in the
XX PT phosphorylated and un-phosphorylated state.
XX PS Claim 27; Page; 65pp; English.
XX CC This sequence represents a green fluorescent protein variant from
XX CC Aequorea victoria in which a wild-type Ser residue at position 65 is
XX CC replaced with a Thr. This variant is used in a method to determine if a
XX CC sample contains protein kinase (PK) activity. The method involves
XX CC contacting the sample with a phosphate donor (PD) and a fluorescent
XX CC protein (FP) substrate for a PK, the protein substrate comprising a FP
XX CC moiety and a phosphorylation site for a PK, where the protein substrate
XX CC exhibits a different fluorescent property in the phosphorylated state
XX CC than in the un-phosphorylated state. The protein substrate is then
XX CC excited and the amount of a fluorescent property that differs in the un-
XX CC phosphorylated state and phosphorylated state is measured, whereby an
XX CC amount that is consistent with the presence of the protein substrate in
XX CC its phosphorylated state indicates the presence of PK activity. The
XX CC method and products can be used in drug screening. They can be used for
XX CC screening for compounds which affect cellular events, including receptor-
XX CC ligand binding, protein-protein interactions or kinase activation, which
XX CC signal to the target kinase. NOTE: This sequence does not appear in the
XX CC specification and has been constructed from the wild-type sequence

```


CC isolated from *Aequorea victoria*. This protein is used in the construction
CC of a mutant protein which emits fluorescence 6 times higher than the
CC conventional fluorescent proteins. The nucleic acid can be used for the
CC recombinant production of the fluorescent protein and also for the
CC determination of activity of a promoter to be tested, comprising steps of
CC introducing the above vector to a host cell and detecting the
CC fluorescence emitted from the cell. They can additionally be used for
CC detecting the targeting activity in the cell of an amino acid sequence to
CC be tested in which the above fluorescent protein is introduced to a cell
CC and the distribution of the fluorescent protein in the cell is observed
XX
SQ Sequence 238 AA;

Query Match 100.0%; Score 1260; DB 2; Length 238;
Best Local Similarity 98.3%; Pred. No. 1.5e-119;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVDNGHKFSVSGEGDATTGKTLTKFCTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVDNGHKFSVSGEGDATTGKTLTKFCTTGKLPVWPPTL 60
QY 61 VTTXXYGVQCFSRYPDHMKRHDFFKSAMPEGYVOERTIFFKDDGNVKTAEVKEFGDTLV 120
DB 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVOERTIFFKDDGNVKTAEVKEFGDTLV 120
QY 121 NRIELKGIDPKEDGNILGHKLEYNNSHNYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180
DB 121 NRIELKGIDPKEDGNILGHKLEYNNSHNYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLXFVTAAGITHGMDELYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLXFVTAAGITHGMDELYK 238

RESULT 13

AAB73552 standard; protein; 238 AA.

AC AAB73552;
DT 07-AUG-2001 (first entry)

DE Wild-type green fluorescent protein (wtGFP).

XX Phenotype selection; non-selectable; fusion protein; stable expression;
KW selectable marker; antibiotic resistance gene; *Escherichia coli*;
KM green fluorescent protein; GFP; wtGFP; pgFP.

OS *Aequorea victoria*.

PN W0200129225-A1.

PD 26-APR-2001.

PF 29-MAR-2000; 2000MO-US008477.

PR 21-OCT-1999; 990US-0160461P.

PR 22-FEB-2000; 2000US-00510097.

PA (PANO-) PANORAMA RES INC.

PI Balint RF;

DR WPI; 2001-282162/29.

DR N-PSDB; AAH20245.

PT Obtaining cells expressing mutant protein, comprises selecting from cells
PT transformed with library of mutagenized protein coding sequences joined
PT to selector protein, which confers growth under selective conditions.

PS Example 2; Page: 52pp; English.

XX The invention relates to methods whereby variants of proteins which do

CC not confer selectable phenotypes can be selected for stable expression in
CC host cells (especially *Escherichia coli*). The methods can be used to
CC obtain mutants of a desired protein optimised for expression in the host
CC cells, for obtaining a mutant protein with enhanced stability relative to
CC the wild-type protein. The methods all involve expressing the protein of
CC interest as a fusion with a protein that can confer a selectable
CC phenotype, such as an antibiotic resistance protein. The transformed host
CC cells are then grown under selective pressure (e.g., presence of
CC antibiotic). The cells able to grow under such conditions are those which
CC contain fusion proteins which are optimised for expression or which are
CC more stable, as this property will also correlate with an increased
CC amount or rate of synthesis of the selectable marker. The invention also
CC discloses mutants of green fluorescent protein (GFP) selected for
CC increased stability using the method of the invention. The present
CC sequence represents wild-type GFP (wtGFP) as encoded by plasmid pgFP
CC (Genbank U17997), which was used in an exemplification of the invention.
CC The present sequence is not shown in the specification, but was derived
CC from Genbank accession number U17997
XX

SQ Sequence 238 AA;

Query Match 100.0%; Score 1260; DB 4; Length 238;
Best Local Similarity 98.3%; Pred. No. 1.5e-119;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVDNGHKFSVSGEGDATTGKTLTKFCTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVDNGHKFSVSGEGDATTGKTLTKFCTTGKLPVWPPTL 60
QY 61 VTTXXYGVQCFSRYPDHMKRHDFFKSAMPEGYVOERTIFFKDDGNVKTAEVKEFGDTLV 120
DB 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVOERTIFFKDDGNVKTAEVKEFGDTLV 120
QY 121 NRIELKGIDPKEDGNILGHKLEYNNSHNYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180
DB 121 NRIELKGIDPKEDGNILGHKLEYNNSHNYIMADKQKNGIKVNFKIRHNIEDGSVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLXFVTAAGITHGMDELYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLXFVTAAGITHGMDELYK 238

RESULT 14

AAE16046 standard; protein; 238 AA.

AC AAE16046;

DT 26-MAR-2002 (first entry)

DE *Aequorea victoria* GFP mutant protein (S65V).

KW Green fluorescent protein; cell lineage tracer; protein localisation;

KW GFP; fusion tag; gene expression marker; fluorescent energy acceptor;

KW immunosay; hybridisation assay; fluorescent energy donor; biosensor;

KW FRET; fluorescence resonance energy transfer; mutant; mutcin.

OS *Aequorea victoria*.

PN Synthetic.

PI Key Location/Qualifiers

FT Misc-difference 65 /note= "Wild type Ser substituted with Val"

PD 29-NOV-2001.

PF 17-MAY-2001; 2001MO-US016149.

PR 19-MAY-2000; 2000US-00575847.

PA (UYOR-) UNIV OREGON STATE.
 XX
 XX Wachter R, Remington SJ;
 XX
 XX WPI, 2002-083084/11.
 XX
 XX
 XX New long wavelength engineered fluorescent proteins, useful as markers
 PT for gene expression, tracers of cell lineage or as fusion tags to monitor
 PT protein localization, or in detection assays, e.g. immunoassays or
 PT hybridization assays.
 XX
 XX
 XX Claim 2; Page; 181pp; English.
 XX
 XX The patent discloses functional engineered fluorescent protein and its
 CC corresponding polynucleotide. The amino acid sequence of the engineered
 CC protein is identical to Aequorea green fluorescent protein (GFP). The
 CC engineered fluorescent proteins of the invention have varied fluorescent
 CC properties and have the ability to respond to ion concentrations via a
 CC change in fluorescent characteristics. They are useful as markers for
 CC gene expression, tracers of cell lineage or as fusion tags to monitor
 CC protein localization within living cells. The engineered fluorescent
 CC proteins are particularly useful for coupling engineered fluorescent
 CC proteins to antibodies, nucleic acids or other receptors for use in
 CC detection assays, e.g. immunoassays or hybridization assays. They are
 CC useful for tracking the movement of proteins in cells or in systems for
 CC detecting induction of transcription and for the simultaneous measurement
 CC of two or more processes within cells. Proteins of the invention are also
 CC useful as fluorescent energy donors or acceptors as well as biosensors
 CC for detecting anions. They are also useful in fluorescence resonance
 CC energy transfer (FRET). The crystal structure of the GFP is useful for
 CC designing mutants having altered fluorescent characteristics which are
 CC particularly useful to identify amino acids whose substitution alters
 CC fluorescent properties of the protein. The crystal structure of the GFP
 CC is also useful for designing mutants having altered anion binding
 CC characteristics which are particularly useful for identifying amino acids
 CC whose substitution alters the specificity and affinity of the binding
 CC site to various anions, and for monitoring anion binding and therefore
 CC the concentration of the anion. The present sequence is Aequorea victoria
 CC GFP mutant protein (S65V). Note: This sequence is not shown in the
 CC specification but is derived from Aequorea victoria wild type green
 CC fluorescent protein shown in figure 3 of the specification (AAE16038)
 CC
 XX
 XX Sequence 238 AA;
 SO
 Query Match 100.0%; Score 1260; DB 5; Length 238;
 Best Local Similarity 98.3%; Pred. No. 1.5e-119;
 Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSKGEELFTGVVPILVELDGDVNGHKFVS GEGEGDATYGLTKLFICTTGLPVPWPTL 60
 DB 1 MSKGEELFTGVVPILVELDGDVNGHKFVS GEGEGDATYGLTKLFICTTGLPVPWPTL 60
 QY 61 VTTXGVVOCFGRYPDMKRHPFKSAMPEGYQERTIFFKDKGNYKTRAEVFEEDTLV 120
 DB 61 VTTFYGVGVCFSRYPDMKRHPFKSAMPEGYQERTIFFKDKGNYKTRAEVFEEDTLV 120
 QY 121 NRLEKGLDIPKEDGNILGHKLENYNSHVVYIMADKQKNGIKRNFIRNIEGVSQVLAD 180
 DB 121 NRLEKGLDIPKEDGNILGHKLENYNSHVVYIMADKQKNGIKRNFIRNIEGVSQVLAD 180
 QY 181 HYQONTPIDGDVPLPDDNHYLSTQSALSKDPNEKRDHMLTKPVTAAAGTTHGDELYK 238
 DB 181 HYQONTPIDGDVPLPDDNHYLSTQSALSKDPNEKRDHMLTKPVTAAAGTTHGDELYK 238
 RESULT 15
 AAE16042
 ID AAE16042 standard; protein; 238 AA.
 XX
 XX AAE16042;
 AC
 XX
 XX 26-MAR-2002 (first entry)
 DT
 XX

DE Aequorea victoria GFP mutant protein (S65T).
 XX
 XX Green fluorescent protein; cell lineage tracer; protein localization;
 KW GFP; fusion tag; gene expression marker; fluorescent energy acceptor;
 KW immunosassay; hybridisation assay; fluorescent energy donor; biosensor;
 KW FRET; fluorescence resonance energy transfer; mutant; mutcin.
 XX
 XX Aequorea victoria.
 OS Synthetic.
 OS
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 65
 FT /note= "Wild type Ser substituted with Thr"
 XX
 XX WO200130147-A2.
 XX
 XX 29-NOV-2001.
 XX
 XX 17-MAY-2001; 2001WO-US016149.
 XX
 XX 19-MAY-2000; 2000US-00575847.
 XX
 XX (UYOR-) UNIV OREGON STATE.
 XX
 XX Wachter R, Remington SJ;
 XX
 XX WPI, 2002-083084/11.
 XX
 XX
 XX New long wavelength engineered fluorescent proteins, useful as markers
 PT for gene expression, tracers of cell lineage or as fusion tags to monitor
 PT protein localization, or in detection assay, e.g. immunoassays or
 PT hybridization assays.
 XX
 XX
 XX Claim 2; Page; 181pp; English.
 XX
 XX The patent discloses functional engineered fluorescent protein and its
 CC corresponding polynucleotide. The amino acid sequence of the engineered
 CC protein is identical to Aequorea green fluorescent protein (GFP). The
 CC engineered fluorescent proteins of the invention have varied fluorescent
 CC properties and have the ability to respond to ion concentrations via a
 CC change in fluorescent characteristics. They are useful as markers for
 CC gene expression, tracers of cell lineage or as fusion tags to monitor
 CC protein localization within living cells. The engineered fluorescent
 CC proteins are particularly useful for coupling engineered fluorescent
 CC proteins to antibodies, nucleic acids or other receptors for use in
 CC detection assays, e.g. immunoassays or hybridization assays. They are
 CC useful for tracking the movement of proteins in cells or in systems for
 CC detecting induction of transcription and for the simultaneous measurement
 CC of two or more processes within cells. Proteins of the invention are also
 CC useful as fluorescent energy donors or acceptors as well as biosensors
 CC for detecting anions. They are also useful in fluorescence resonance
 CC energy transfer (FRET). The crystal structure of the GFP is useful for
 CC designing mutants having altered fluorescent characteristics which are
 CC particularly useful to identify amino acids whose substitution alters
 CC fluorescent properties of the protein. The crystal structure of the GFP
 CC is also useful for designing mutants having altered anion binding
 CC characteristics which are particularly useful for identifying amino acids
 CC whose substitution alters the specificity and affinity of the binding
 CC site to various anions, and for monitoring anion binding and therefore
 CC the concentration of the anion. The present sequence is Aequorea victoria
 CC GFP mutant protein (S65V). Note: This sequence is not shown in the
 CC specification but is derived from Aequorea victoria wild type green
 CC fluorescent protein shown in figure 3 of the specification (AAE16038)
 CC
 XX
 XX Sequence 238 AA;
 SO
 Query Match 100.0%; Score 1260; DB 5; Length 238;
 Best Local Similarity 98.3%; Pred. No. 1.5e-119;
 Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSKGEELFTGVVPILVELDGDVNGHKFVS GEGEGDATYGLTKLFICTTGLPVPWPTL 60
 DB 1 MSKGEELFTGVVPILVELDGDVNGHKFVS GEGEGDATYGLTKLFICTTGLPVPWPTL 60

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 28, 2005, 14:21:34 ; Search time 43 Seconds
(without alignment)
413.174 Million cell updates/sec

Title: US-10-757-624-2-COPY

Perfect score: 1260
Sequence: 1 MSKGBELFTGVVPIVLELDG.....VLKRFVTAAGITTHGMDELYK 238

Scoring table: BLOSUM62DX

Gapop 10.0, Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/ECTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 1260 | 100.0 | 238 | 1 | US-08-753-143-2 |
| 2 | 1260 | 100.0 | 238 | 2 | US-08-679-865-2 |
| 3 | 1260 | 100.0 | 238 | 2 | US-08-680-876-2 |
| 4 | 1260 | 100.0 | 238 | 2 | US-08-792-553-2 |
| 5 | 1260 | 100.0 | 238 | 3 | US-08-753-144-2 |
| 6 | 1260 | 100.0 | 238 | 3 | US-09-094-359-2 |
| 7 | 1260 | 100.0 | 238 | 3 | US-09-172-063-2 |
| 8 | 1260 | 100.0 | 238 | 3 | US-09-263-975-2 |
| 9 | 1260 | 100.0 | 238 | 3 | US-08-727-452-2 |
| 10 | 1260 | 100.0 | 238 | 4 | US-09-418-785-1 |
| 11 | 1260 | 100.0 | 238 | 4 | US-09-129-192C-2 |
| 12 | 1260 | 100.0 | 238 | 4 | US-09-129-192C-74 |
| 13 | 1260 | 100.0 | 238 | 4 | US-09-602-641-2 |
| 14 | 1260 | 100.0 | 238 | 4 | US-10-024-686A-2 |
| 15 | 1260 | 100.0 | 238 | 4 | US-10-024-686A-2 |
| 16 | 1260 | 100.0 | 238 | 4 | US-08-594-575C-2 |
| 17 | 1260 | 99.7 | 238 | 1 | US-08-337-915A-2 |
| 18 | 1256 | 99.7 | 238 | 3 | US-08-893-327-16 |
| 19 | 1256 | 99.7 | 238 | 3 | US-09-121-539-1 |
| 20 | 1256 | 99.7 | 238 | 4 | US-09-214-909-2 |
| 21 | 1256 | 99.7 | 238 | 4 | US-09-479-645A-10 |
| 22 | 1256 | 99.7 | 238 | 4 | US-09-479-645A-159 |
| 23 | 1256 | 99.7 | 238 | 4 | US-09-472-065A-2 |
| 24 | 1256 | 99.7 | 238 | 4 | US-09-472-065A-4 |
| 25 | 1256 | 99.7 | 238 | 4 | US-09-472-065A-5 |
| 26 | 1256 | 99.7 | 238 | 4 | US-09-472-065A-6 |
| 27 | 1256 | 99.7 | 238 | 4 | US-09-920-922-4 |

| | | | | | | |
|----|------|------|------|---|--------------------|--------------------|
| 28 | 1256 | 99.7 | 238 | 5 | PCT-US95-14692-2 | Sequence 2, Appl1 |
| 29 | 1256 | 99.7 | 247 | 3 | US-08-893-327-18 | Sequence 18, Appl1 |
| 30 | 1256 | 99.7 | 1070 | 4 | US-09-091-042A-2 | Sequence 2, Appl1 |
| 31 | 1256 | 99.7 | 1452 | 3 | US-09-127-227-2 | Sequence 2, Appl1 |
| 32 | 1255 | 99.6 | 238 | 4 | US-09-023-946B-35 | Sequence 35, Appl1 |
| 33 | 1255 | 99.6 | 239 | 4 | US-08-646-538-2 | Sequence 2, Appl1 |
| 34 | 1255 | 99.6 | 239 | 3 | US-09-503-222-2 | Sequence 2, Appl1 |
| 35 | 1253 | 99.4 | 238 | 4 | US-09-603-448-7 | Sequence 7, Appl1 |
| 36 | 1253 | 99.4 | 238 | 4 | US-09-603-448-24 | Sequence 24, Appl1 |
| 37 | 1253 | 99.4 | 243 | 4 | US-09-479-645A-94 | Sequence 94, Appl1 |
| 38 | 1253 | 99.4 | 243 | 4 | US-09-479-645A-96 | Sequence 96, Appl1 |
| 39 | 1253 | 99.4 | 243 | 4 | US-09-479-645A-98 | Sequence 98, Appl1 |
| 40 | 1253 | 99.4 | 243 | 4 | US-09-479-645A-100 | Sequence 100, App |
| 41 | 1253 | 99.4 | 243 | 4 | US-09-479-645A-102 | Sequence 102, App |
| 42 | 1253 | 99.4 | 243 | 4 | US-09-479-645A-104 | Sequence 104, App |
| 43 | 1253 | 99.4 | 243 | 4 | US-09-479-645A-110 | Sequence 110, App |
| 44 | 1252 | 99.4 | 243 | 4 | US-09-479-645A-88 | Sequence 88, Appl1 |
| 45 | 1252 | 99.4 | 243 | 4 | US-09-479-645A-90 | Sequence 90, Appl1 |

ALIGNMENTS

```
RESULT 1
US-08-753-143-2
; Sequence 2, Application US/08753143A
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; TITLE OF INVENTION: MODIFIED GREEN FLUORESCENT PROTEINS
; FILE REFERENCE: 07257/032003
; CURRENT APPLICATION NUMBER: US/08/753,143A
; CURRENT FILING DATE: 1996-11-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-08-753-143-2

Query Match      100.0%; Score 1260; DB 1; Length 238;
Best Local Similarity 98.3%; Pred. NO. 5.6e-124;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      1 MSKGBELFTGVVPIVLELDGDNHGKFSVSGEGEDATYGKLTLLKFTCTGKLPVWPPTL 60
DB      1 MSKGBELFTGVVPIVLELDGDNHGKFSVSGEGEDATYGKLTLLKFTCTGKLPVWPPTL 60
QY      61 VTTXXYGVOCFSRYPDHHKRRDPFKSAMPBEGVQERTTFFKODGNYKTRAEYKFGSDTLV 120
DB      61 VTFYGVQCSRRYPDHHKRRDPFKSAMPBEGVQERTTFFKODGNYKTRAEYKFGSDTLV 120
QY      121 NRLEKIDGDFKEDGNILHKLKLEYNNSHVVYIMADKONGIKVNFKIRHNIEDGXVOLAD 180
DB      121 NRLEKIDGDFKEDGNILHKLKLEYNNSHVVYIMADKONGIKVNFKIRHNIEDGXVOLAD 180
QY      181 HYQNTPIGDEGVLLPDDNYLSTGSALSXPNEKRDHVLKFPVTAAGITTHGMDELYK 238
DB      181 HYQNTPIGDEGVLLPDDNYLSTGSALSXPNEKRDHVLKFPVTAAGITTHGMDELYK 238

RESULT 2
US-08-679-865-2
; Sequence 2, Application US/08679865
; Patent No. 5912137
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Assays for Protein Kinases Using
; TITLE OF INVENTION: Fluorescent Protein Substrates
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/679,865
FILING DATE: 16-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Storella, John S.
REGISTRATION NUMBER: 32,944
REFERENCE/DOCKET NUMBER: 023072-069000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 238 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-679-865-2

Query Match 100.0%; Score 1260; DB 2; Length 238;
Best Local Similarity 98.3%; Pred. No. 5.6e-124;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVNGHKFSVSGEGDATYGLTKLFICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVNGHKFSVSGEGDATYGLTKLFICTTGKLPVWPPTL 60
QY 61 VTTXXYGVOCFSRYPDHMKRHDFFKSAPEGYOERTIFFKDDGNKTRAEVFEEDTLV 120
DB 61 VTTFSYGVOCFSRYPDHMKRHDFFKSAPEGYOERTIFFKDDGNKTRAEVFEEDTLV 120
QY 121 NRTELKIDFKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGVQVLAD 180
DB 121 NRTELKIDFKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGVQVLAD 180
QY 181 HYQONTPIGDPVLLPDNHYLSTQSALSKDPNEKRDHNVLLXFVTAAGITHGMDELYK 238
DB 181 HYQONTPIGDPVLLPDNHYLSTQSALSKDPNEKRDHNVLLXFVTAAGITHGMDELYK 238

RESULT 3
US-08-680-876-2
Sequence 2, Application US/08680876
Patent No. 592558
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
APPLICANT: Cubitt, Andrew B.
TITLE OF INVENTION: Assays for Protein Kinases Using
TITLE OF INVENTION: Fluorescent Protein Substrates
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/680,876
FILING DATE: 16-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Storella, John S.
REGISTRATION NUMBER: 32,944
REFERENCE/DOCKET NUMBER: 023072-069200
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 238 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-680-876-2

Query Match 100.0%; Score 1260; DB 2; Length 238;
Best Local Similarity 98.3%; Pred. No. 5.6e-124;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVNGHKFSVSGEGDATYGLTKLFICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVNGHKFSVSGEGDATYGLTKLFICTTGKLPVWPPTL 60
QY 61 VTTXXYGVOCFSRYPDHMKRHDFFKSAPEGYOERTIFFKDDGNKTRAEVFEEDTLV 120
DB 61 VTTFSYGVOCFSRYPDHMKRHDFFKSAPEGYOERTIFFKDDGNKTRAEVFEEDTLV 120
QY 121 NRTELKIDFKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGVQVLAD 180
DB 121 NRTELKIDFKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGVQVLAD 180
QY 181 HYQONTPIGDPVLLPDNHYLSTQSALSKDPNEKRDHNVLLXFVTAAGITHGMDELYK 238
DB 181 HYQONTPIGDPVLLPDNHYLSTQSALSKDPNEKRDHNVLLXFVTAAGITHGMDELYK 238

RESULT 4
US-08-792-553-2
Sequence 2, Application US/08792553
Patent No. 5981200
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
APPLICANT: Heim, Roger
TITLE OF INVENTION: Tandem Fluorescent Protein Constructs
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: California
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/792,553
FILING DATE: 31-JAN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lisa A. Hallie, Ph.D.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07257/041001/UC 96-160-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-678-5070
TELEFAX: 619-678-5099
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 238 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-792-553-2

Query Match 100.0%; Score 1260; DB 2; Length 238;
Best Local Similarity 98.3%; Pred. No. 5,6e-124;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVDNGHKFSVSGEGEDATYGLTLKFCITGKLPVPWPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVDNGHKFSVSGEGEDATYGLTLKFCITGKLPVPWPTL 60
QY 61 VTTXXYGVQCFSRYPDHNKRHDFPKSAMPEGVOERTIFFKDDGNYKTRAAYKFGDITLV 120
DB 61 VTTFSYGVQCFSRYPDHNKRHDFPKSAMPEGVOERTIFFKDDGNYKTRAAYKFGDITLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVTIMADKQNGIKVNFKIRHNIEDGVQLAD 180
DB 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVTIMADKQNGIKVNFKIRHNIEDGVQLAD 180
QY 181 HYQONTPIGDGFVLLPDNHYLSTQSALSKDPNEKRDHVLXFTVAAGITTHGMDELYK 238
DB 181 HYQONTPIGDGFVLLPDNHYLSTQSALSKDPNEKRDHVLXFTVAAGITTHGMDELYK 238

RESULT 5
US-08-753-144-2
Sequence 2, Application US/08753144
Patent No. 6066476

GENERAL INFORMATION:
APPLICANT: Taisen, Roger Y.

APPLICANT: Heim, Roger
TITLE OF INVENTION: MODIFIED GREEN FLUORESCENT PROTEINS
NUMBER OF SEQUENCES: 5
CURRENT FILING DATE: 1998-06-09

ADDRESS: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA

ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/753,144
FILING DATE: 20-NOV-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/727,452
FILING DATE: 10-OCT-1996
APPLICATION NUMBER: US95/14692
FILING DATE: 13-NOV-1995
APPLICATION NUMBER: 08/357,915
FILING DATE: 10-NOV-1994

ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07257/032002
TELECOMMUNICATION INFORMATION:

TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 238 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal

US-08-753-144-2

Query Match 100.0%; Score 1260; DB 3; Length 238;
Best Local Similarity 98.3%; Pred. No. 5,6e-124;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVDNGHKFSVSGEGEDATYGLTLKFCITGKLPVPWPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVDNGHKFSVSGEGEDATYGLTLKFCITGKLPVPWPTL 60
QY 61 VTTXXYGVQCFSRYPDHNKRHDFPKSAMPEGVOERTIFFKDDGNYKTRAAYKFGDITLV 120
DB 61 VTTFSYGVQCFSRYPDHNKRHDFPKSAMPEGVOERTIFFKDDGNYKTRAAYKFGDITLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVTIMADKQNGIKVNFKIRHNIEDGVQLAD 180
DB 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVTIMADKQNGIKVNFKIRHNIEDGVQLAD 180
QY 181 HYQONTPIGDGFVLLPDNHYLSTQSALSKDPNEKRDHVLXFTVAAGITTHGMDELYK 238
DB 181 HYQONTPIGDGFVLLPDNHYLSTQSALSKDPNEKRDHVLXFTVAAGITTHGMDELYK 238

RESULT 6
US-09-094-359-2
Sequence 2, Application US/09094359
Patent No. 6140132

GENERAL INFORMATION:
APPLICANT: Taisen, Roger Y.

APPLICANT: Miyawaki, Atsushi
APPLICANT: Llopis, Juan
TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
FILE REFERENCE: 07257/067001
CURRENT APPLICATION NUMBER: US/09/094,359
CURRENT FILING DATE: 1998-06-09

NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 2
LENGTH: 238

TYPE: PRT
ORGANISM: Aequorea victoria
US-09-094-359-2

Query Match 100.0%; Score 1260; DB 3; Length 238;
Best Local Similarity 98.3%; Pred. No. 5,6e-124;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVDNGHKFSVSGEGEDATYGLTLKFCITGKLPVPWPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVDNGHKFSVSGEGEDATYGLTLKFCITGKLPVPWPTL 60
QY 61 VTTXXYGVQCFSRYPDHNKRHDFPKSAMPEGVOERTIFFKDDGNYKTRAAYKFGDITLV 120
DB 61 VTTFSYGVQCFSRYPDHNKRHDFPKSAMPEGVOERTIFFKDDGNYKTRAAYKFGDITLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVTIMADKQNGIKVNFKIRHNIEDGVQLAD 180
DB 121 NRIELKGIDFKEDGNILGHKLEYNNSHNVTIMADKQNGIKVNFKIRHNIEDGVQLAD 180
QY 181 HYQONTPIGDGFVLLPDNHYLSTQSALSKDPNEKRDHVLXFTVAAGITTHGMDELYK 238
DB 181 HYQONTPIGDGFVLLPDNHYLSTQSALSKDPNEKRDHVLXFTVAAGITTHGMDELYK 238

RESULT 7
US-09-172-063-2
Sequence 2, Application US/09172063
Patent No. 6150176

GENERAL INFORMATION:
APPLICANT: Taisen, Roger Y.
APPLICANT: Miyawaki, Atsushi
APPLICANT: Llopis, Juan

```

; APPLICANT: Machter, Rebekka M.
; APPLICANT: Remington, S. James
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; TITLE OF INVENTION: MEASURING THE PH OF A BIOLOGICAL SAMPLE
; FILE REFERENCE: 07257/071001
; CURRENT APPLICATION NUMBER: US/09/172,063
; CURRENT FILING DATE: 1998-10-13
; EARLIER APPLICATION NUMBER: 09/094,359
; EARLIER FILING DATE: 1998-06-09
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-09-172-063-2

Query Match          100.0%; Score 1260; DB 3; Length 238;
Best Local Similarity 98.3%; Pred. No. 5.6e-124;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVDVNGHKFVSGBEGGDATYGLTLKFICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVDVNGHKFVSGBEGGDATYGLTLKFICTTGKLPVWPPTL 60
QY 61 VTTXXYGVQCFSRYPDHMKRHDFFKSAWPEGYOERTIFFKDDGNYKTRAIVKFEEDTLV 120
DB 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAWPEGYOERTIFFKDDGNYKTRAIVKFEEDTLV 120
QY 121 NRLEKIGIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGVQLAD 180
DB 121 NRLEKIGIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLXFTVTAAGITGMDLYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLXFTVTAAGITGMDLYK 238

RESULT 8
US-09-263-975-2
; Sequence 2, Application US/09263975
; Patent No. 6248550
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Assays for Protein Kinases Using
; TITLE OF INVENTION: Fluorescent Protein Substrates
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/263,975
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/679,865
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Storella, John S.
; REGISTRATION NUMBER: 32,944
; REFERENCE/DOCKET NUMBER: 023072-069000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
```

```

; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-263-975-2

Query Match          100.0%; Score 1260; DB 3; Length 238;
Best Local Similarity 98.3%; Pred. No. 5.6e-124;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVDVNGHKFVSGBEGGDATYGLTLKFICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVDVNGHKFVSGBEGGDATYGLTLKFICTTGKLPVWPPTL 60
QY 61 VTTXXYGVQCFSRYPDHMKRHDFFKSAWPEGYOERTIFFKDDGNYKTRAIVKFEEDTLV 120
DB 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAWPEGYOERTIFFKDDGNYKTRAIVKFEEDTLV 120
QY 121 NRLEKIGIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGVQLAD 180
DB 121 NRLEKIGIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLXFTVTAAGITGMDLYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLXFTVTAAGITGMDLYK 238

RESULT 9
US-08-727-452-2
; Sequence 2, Application US/08727452A
; Patent No. 6319669
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Helm, Roger
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR DETECTION OF ANALYTES
; FILE REFERENCE: 07257/032001
; CURRENT APPLICATION NUMBER: US/08/727,452A
; CURRENT FILING DATE: 1996-03-20
; EARLIER APPLICATION NUMBER: PCT/US95/14692
; EARLIER FILING DATE: 1995-11-13
; EARLIER APPLICATION NUMBER: US 07/337,915
; EARLIER FILING DATE: 1994-11-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-08-727-452-2

Query Match          100.0%; Score 1260; DB 3; Length 238;
Best Local Similarity 98.3%; Pred. No. 5.6e-124;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVDVNGHKFVSGBEGGDATYGLTLKFICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVDVNGHKFVSGBEGGDATYGLTLKFICTTGKLPVWPPTL 60
QY 61 VTTXXYGVQCFSRYPDHMKRHDFFKSAWPEGYOERTIFFKDDGNYKTRAIVKFEEDTLV 120
DB 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAWPEGYOERTIFFKDDGNYKTRAIVKFEEDTLV 120
QY 121 NRLEKIGIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGVQLAD 180
DB 121 NRLEKIGIDFKEDGNILGHKLEYNVNSHNYIMADKQNGIKVNFKIRHNIEDGVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLXFTVTAAGITGMDLYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLXFTVTAAGITGMDLYK 238
```

RESULT 10
US-09-418-785-1
; Sequence 1, Application US/09418785
; Patent No. 6414119
; GENERAL INFORMATION:
; APPLICANT: Fisher, Hugh
; TITLE OF INVENTION: Rapidly Greening, Low Oxygen Mutant of
; TITLE OF INVENTION: the Aequorea victoria Green Fluorescent Protein
; FILE REFERENCE: RUC 99-0011
; CURRENT APPLICATION NUMBER: US/09/418,785
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/104,563
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 238
; TYPE: PRF
; ORGANISM: Aequorea victoria
; PUBLICATION INFORMATION:
; AUTHORS: Presher, D.C. et al.
; TITLE: Primary structure of the Aequorea victoria green-f
; JOURNAL: Gene
; VOLUME: 111
; PAGES: 229-233
; DATE: 1992-01-01
; DATABASE ACCESSION NUMBER: Genbank M62653
; DATABASE ENTRY DATE: 1993-04-26
US-09-418-785-1

Query Match 100.0%; Score 1260; DB 4; Length 238;
Best Local Similarity 98.3%; Pred. No. 5.6e-124;

Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVELDGVNGHKFSVSGEGEDATYGLTKLFICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVELDGVNGHKFSVSGEGEDATYGLTKLFICTTGKLPVWPPTL 60
QY 61 VTTXXYGVQCFSRYPDHNKRHDFPKSAMPEGVQERTIFFKODGNYKTRAEVKFGDITLV 120
DB 61 VTTFSYGVQCFSRYPDHNKRHDFPKSAMPEGVQERTIFFKODGNYKTRAEVKFGDITLV 120
QY 121 NRIELKIDFKEDGNILGHKLEYNNSHNVYIMADKXNGIKVNFKIRNINIEDGSVQLAD 180
DB 121 NRIELKIDFKEDGNILGHKLEYNNSHNVYIMADKXNGIKVNFKIRNINIEDGSVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLKFTVPAAGITTHGMDELYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLKFTVPAAGITTHGMDELYK 238

RESULT 11

US-09-129-192C-2
; Sequence 2, Application US/09129192C
; Patent No. 6495664
; GENERAL INFORMATION:
; APPLICANT: Aurora Biosciences Corporation
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Fluorescent Protein Sensors of Post-Translational Modifications
; FILE REFERENCE: AURO1270 (08366/031001)
; CURRENT APPLICATION NUMBER: US/09/129,192C
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRF
; ORGANISM: Aequorea

US-09-129-192C-2

Query Match 100.0%; Score 1260; DB 4; Length 238;
Best Local Similarity 98.3%; Pred. No. 5.6e-124;

Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSKGEELFTGVVPIVELDGVNGHKFSVSGEGEDATYGLTKLFICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVELDGVNGHKFSVSGEGEDATYGLTKLFICTTGKLPVWPPTL 60
QY 61 VTTXXYGVQCFSRYPDHNKRHDFPKSAMPEGVQERTIFFKODGNYKTRAEVKFGDITLV 120
DB 61 VTTFSYGVQCFSRYPDHNKRHDFPKSAMPEGVQERTIFFKODGNYKTRAEVKFGDITLV 120
QY 121 NRIELKIDFKEDGNILGHKLEYNNSHNVYIMADKXNGIKVNFKIRNINIEDGSVQLAD 180
DB 121 NRIELKIDFKEDGNILGHKLEYNNSHNVYIMADKXNGIKVNFKIRNINIEDGSVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLKFTVPAAGITTHGMDELYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLKFTVPAAGITTHGMDELYK 238

RESULT 12

US-09-129-192C-74

; Sequence 74, Application US/09129192C
; Patent No. 6495664
; GENERAL INFORMATION:
; APPLICANT: Aurora Biosciences Corporation
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Fluorescent Protein Sensors of Post-Translational Modifications
; FILE REFERENCE: AURO1270 (08366/031001)
; CURRENT APPLICATION NUMBER: US/09/129,192C
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 74
; LENGTH: 238
; TYPE: PRF
; ORGANISM: Aequorea green fluorescent protein phosphorylation mutant

US-09-129-192C-74

Query Match 100.0%; Score 1260; DB 4; Length 238;
Best Local Similarity 98.3%; Pred. No. 5.6e-124;

Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVELDGVNGHKFSVSGEGEDATYGLTKLFICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVELDGVNGHKFSVSGEGEDATYGLTKLFICTTGKLPVWPPTL 60
QY 61 VTTXXYGVQCFSRYPDHNKRHDFPKSAMPEGVQERTIFFKODGNYKTRAEVKFGDITLV 120
DB 61 VTTFSYGVQCFSRYPDHNKRHDFPKSAMPEGVQERTIFFKODGNYKTRAEVKFGDITLV 120
QY 121 NRIELKIDFKEDGNILGHKLEYNNSHNVYIMADKXNGIKVNFKIRNINIEDGSVQLAD 180
DB 121 NRIELKIDFKEDGNILGHKLEYNNSHNVYIMADKXNGIKVNFKIRNINIEDGSVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLKFTVPAAGITTHGMDELYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLKFTVPAAGITTHGMDELYK 238

RESULT 13

US-09-602-641-2

; Sequence 2, Application US/09602641
; Patent No. 6608189
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Llopis, Juan
; APPLICANT: Wachter, Rebekka M.
; APPLICANT: Remington, S. James
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; TITLE OF INVENTION: MEASURING THE PH OF A BIOLOGICAL SAMPLE
; FILE REFERENCE: 07257/071001
; CURRENT APPLICATION NUMBER: US/09/602,641

/ CURRENT FILING DATE: 2000-06-22
/ PRIOR APPLICATION NUMBER: 09/172,063
/ PRIOR FILING DATE: 1998-10-13
/ NUMBER OF SEQ ID NOS: 38
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 238
/ TYPE: PRT
/ ORGANISM: Aequorea victoria
US-09-602-641-2

Query Match 100.0%; Score 1260; DB 4; Length 238;
Best Local Similarity 98.3%; Pred. No. 5,6e-124;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGDATYGLTLKFICTTGKLPVPMPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGDATYGLTLKFICTTGKLPVPMPTL 60
QY 61 VTTXXYGVQCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNKYTRAEVKFEEDTLV 120
DB 61 VTTXXYGVQCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNKYTRAEVKFEEDTLV 120
QY 121 NRTELKIDFKEDGNILGHKLEYNNSHNYIMADKQKGIKVNFKIRHNIEDGSVQLAD 180
DB 121 NRTELKIDFKEDGNILGHKLEYNNSHNYIMADKQKGIKVNFKIRHNIEDGSVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLFVTAAGITTHGMDLYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLFVTAAGITTHGMDLYK 238

RESULT 14
US-09-704-463-2
/ Sequence 2, Application US/09704463
/ Patent No. 6627449
/ GENERAL INFORMATION:
/ APPLICANT: Tsien, Roger Y.
/ APPLICANT: Miyawaki, Akiuehi
/ APPLICANT: Liopis, Juan
/ TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
/ TITLE OF INVENTION: MEASURING THE PH OF A BIOLOGICAL SAMPLE
/ FILE REFERENCE: 07257/067001
/ CURRENT APPLICATION NUMBER: US/09/704,463
/ CURRENT FILING DATE: 2000-10-31
/ PRIOR APPLICATION NUMBER: 09/094,359
/ PRIOR FILING DATE: 1998-06-09
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 238
/ TYPE: PRT
/ ORGANISM: Aequorea victoria
US-09-704-463-2

Query Match 100.0%; Score 1260; DB 4; Length 238;
Best Local Similarity 98.3%; Pred. No. 5,6e-124;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGDATYGLTLKFICTTGKLPVPMPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGDATYGLTLKFICTTGKLPVPMPTL 60
QY 61 VTTXXYGVQCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNKYTRAEVKFEEDTLV 120
DB 61 VTTXXYGVQCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNKYTRAEVKFEEDTLV 120
QY 121 NRTELKIDFKEDGNILGHKLEYNNSHNYIMADKQKGIKVNFKIRHNIEDGSVQLAD 180
DB 121 NRTELKIDFKEDGNILGHKLEYNNSHNYIMADKQKGIKVNFKIRHNIEDGSVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLFVTAAGITTHGMDLYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLFVTAAGITTHGMDLYK 238

DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLFVTAAGITTHGMDLYK 238
RESULT 15
US-10-024-686A-2
/ Sequence 2, Application US/10024686A
/ Patent No. 6800733
/ GENERAL INFORMATION:
/ APPLICANT: Tsien, Roger Y.
/ APPLICANT: Heim, Roger
/ TITLE OF INVENTION: Modified Green Fluorescent Proteins
/ FILE REFERENCE: 39754/0861
/ CURRENT APPLICATION NUMBER: US/10/024,686A
/ CURRENT FILING DATE: 2001-12-17
/ PRIOR APPLICATION NUMBER: 09/057,995
/ PRIOR FILING DATE: 1998-04-09
/ PRIOR APPLICATION NUMBER: 08/753,144
/ PRIOR FILING DATE: 1996-11-20
/ PRIOR APPLICATION NUMBER: 08/727,452
/ PRIOR FILING DATE: 1997-03-20
/ PRIOR APPLICATION NUMBER: PCT/US95/14692
/ PRIOR FILING DATE: 1995-11-13
/ PRIOR APPLICATION NUMBER: 08/337,915
/ PRIOR FILING DATE: 1994-11-10
/ NUMBER OF SEQ ID NOS: 5
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 238
/ TYPE: PRT
/ ORGANISM: Aequorea victoria
US-10-024-686A-2

Query Match 100.0%; Score 1260; DB 4; Length 238;
Best Local Similarity 98.3%; Pred. No. 5,6e-124;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGDATYGLTLKFICTTGKLPVPMPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGDATYGLTLKFICTTGKLPVPMPTL 60
QY 61 VTTXXYGVQCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNKYTRAEVKFEEDTLV 120
DB 61 VTTXXYGVQCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNKYTRAEVKFEEDTLV 120
QY 121 NRTELKIDFKEDGNILGHKLEYNNSHNYIMADKQKGIKVNFKIRHNIEDGSVQLAD 180
DB 121 NRTELKIDFKEDGNILGHKLEYNNSHNYIMADKQKGIKVNFKIRHNIEDGSVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLFVTAAGITTHGMDLYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLFVTAAGITTHGMDLYK 238

Search completed: February 28, 2005, 14:31.07
Job time : 45 secs

; INFORMATION FOR SEQ ID NO: 2;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-884-681-2

Query Match 100.0%; Score 1260; DB 9; Length 238;
Best Local Similarity 98.3%; Pred. No. 3e-112;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGELFTGVVPIVLELDGVNGHKFSVSGEGGDATYGLTLKFCITGKLPVWPPTL 60
Db 1 MSKGELFTGVVPIVLELDGVNGHKFSVSGEGGDATYGLTLKFCITGKLPVWPPTL 60
QY 61 VTTXXYGVQCFSRYPDHMKRHDFFKSAWPEGVQERTIFFKDGNYKTRAEVKFEEDTLV 120
Db 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAWPEGVQERTIFFKDGNYKTRAEVKFEEDTLV 120
QY 121 NRTELKIDPFKEDGNILGHKLEYNVNSHNVYIMADKXNGIKVNFKIRHNIEDGVQLAD 180
Db 121 NRTELKIDPFKEDGNILGHKLEYNVNSHNVYIMADKXNGIKVNFKIRHNIEDGVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLXFTVAAGITTHGMDELYK 238
Db 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLXFTVAAGITTHGMDELYK 238

RESULT 2

US-09-967-301-2
; Sequence 2, Application US/09967301
; Publication No. US20030175859A1
; GENERAL INFORMATION:
; APPLICANT: Stubbs, Simon L.
; APPLICANT: Jones, Anne E.
; APPLICANT: Michael, Nigel P.
; APPLICANT: Thomas, Nicholas
; TITLE OF INVENTION: Fluorescent Proteins
; FILE REFERENCE: PA0111
; CURRENT APPLICATION NUMBER: US/09/967,301
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: GB 0109858.1
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-09-967-301-2

Query Match 100.0%; Score 1260; DB 10; Length 238;
Best Local Similarity 98.3%; Pred. No. 3e-112;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGELFTGVVPIVLELDGVNGHKFSVSGEGGDATYGLTLKFCITGKLPVWPPTL 60
Db 1 MSKGELFTGVVPIVLELDGVNGHKFSVSGEGGDATYGLTLKFCITGKLPVWPPTL 60
QY 61 VTTXXYGVQCFSRYPDHMKRHDFFKSAWPEGVQERTIFFKDGNYKTRAEVKFEEDTLV 120
Db 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAWPEGVQERTIFFKDGNYKTRAEVKFEEDTLV 120
QY 121 NRTELKIDPFKEDGNILGHKLEYNVNSHNVYIMADKXNGIKVNFKIRHNIEDGVQLAD 180
Db 121 NRTELKIDPFKEDGNILGHKLEYNVNSHNVYIMADKXNGIKVNFKIRHNIEDGVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLXFTVAAGITTHGMDELYK 238
Db 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLXFTVAAGITTHGMDELYK 238

RESULT 3

US-09-967-301-3
; Sequence 3, Application US/09967301
; Publication No. US20030175859A1
; GENERAL INFORMATION:
; APPLICANT: Stubbs, Simon L.
; APPLICANT: Jones, Anne E.
; APPLICANT: Michael, Nigel P.
; APPLICANT: Thomas, Nicholas
; TITLE OF INVENTION: Fluorescent Proteins
; FILE REFERENCE: PA0111
; CURRENT APPLICATION NUMBER: US/09/967,301
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: GB 0109858.1
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-967-301-3

Query Match 100.0%; Score 1260; DB 10; Length 238;
Best Local Similarity 98.3%; Pred. No. 3e-112;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGELFTGVVPIVLELDGVNGHKFSVSGEGGDATYGLTLKFCITGKLPVWPPTL 60
Db 1 MSKGELFTGVVPIVLELDGVNGHKFSVSGEGGDATYGLTLKFCITGKLPVWPPTL 60
QY 61 VTTXXYGVQCFSRYPDHMKRHDFFKSAWPEGVQERTIFFKDGNYKTRAEVKFEEDTLV 120
Db 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAWPEGVQERTIFFKDGNYKTRAEVKFEEDTLV 120
QY 121 NRTELKIDPFKEDGNILGHKLEYNVNSHNVYIMADKXNGIKVNFKIRHNIEDGVQLAD 180
Db 121 NRTELKIDPFKEDGNILGHKLEYNVNSHNVYIMADKXNGIKVNFKIRHNIEDGVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLXFTVAAGITTHGMDELYK 238
Db 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLXFTVAAGITTHGMDELYK 238

RESULT 4

US-09-967-301-4
; Sequence 4, Application US/09967301
; Publication No. US20030175859A1
; GENERAL INFORMATION:
; APPLICANT: Stubbs, Simon L.
; APPLICANT: Jones, Anne E.
; APPLICANT: Michael, Nigel P.
; APPLICANT: Thomas, Nicholas
; TITLE OF INVENTION: Fluorescent Proteins
; FILE REFERENCE: PA0111
; CURRENT APPLICATION NUMBER: US/09/967,301
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: GB 0109858.1
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-967-301-4

Query Match 100.0%; Score 1260; DB 10; Length 238;
Best Local Similarity 98.3%; Pred. No. 3e-112;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPLIVELDGVNGHKFSVSGEGDATTGKLTLLKFTCTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPLIVELDGVNGHKFSVSGEGDATTGKLTLLKFTCTTGKLPVWPPTL 60
QY 61 VTTXXYGVQCSRRYPDHMKRHDFFKSAMPEGVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
DB 61 VTTLYGVQCSRRYPDHMKRHDFFKSAMPEGVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKIDPFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDGVSQQLAD 180
DB 121 NRIELKIDPFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDGVSQQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHNVLLFVYTAAGITTHGMDELYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHNVLLFVYTAAGITTHGMDELYK 238

RESULT 5

US-10-024-686-2
; Sequence 2, Application US/10024686
; Publication No. US20020123113A1
; GENERAL INFORMATION:

APPLICANT: Tsien, Roger Y.

Heim, Roger

TITLE OF INVENTION: MODIFIED GREEN FLUORESCENT PROTEINS

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 4225 Executive Square, Suite 1400

CITY: La Jolla

STATE: CA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows 95

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/024, 686

FILING DATE: 17-Dec-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/057, 995

FILING DATE: <Unknown>

APPLICATION NUMBER: 08/727, 452

APPLICATION NUMBER: US95/14692

FILING DATE: 13-NOV-1995

APPLICATION NUMBER: 08/337, 915

FILING DATE: 10-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: Haile, Lisa A.

REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: 07257/032002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619/678-5070

TELEFAX: 619/678-5099

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 238 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: internal

SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-024-686-2

Query Match 100.0%; Score 1260; DB 13; Length 238;

Best Local Similarity 98.3%; Pred. No. 3e-112;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPLIVELDGVNGHKFSVSGEGDATTGKLTLLKFTCTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPLIVELDGVNGHKFSVSGEGDATTGKLTLLKFTCTTGKLPVWPPTL 60
QY 61 VTTXXYGVQCSRRYPDHMKRHDFFKSAMPEGVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
DB 61 VTTBSYGVQCSRRYPDHMKRHDFFKSAMPEGVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKIDPFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDGVSQQLAD 180
DB 121 NRIELKIDPFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDGVSQQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHNVLLFVYTAAGITTHGMDELYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHNVLLFVYTAAGITTHGMDELYK 238

RESULT 6

US-10-057-505-2
; Sequence 2, Application US/10057505
; Publication No. US20020164674A1
; GENERAL INFORMATION:

APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA

APPLICANT: AURORA BIOSCIENCES CORPORATION

APPLICANT: TSJEN, Roger

APPLICANT: HEIM, Roger

APPLICANT: CUBITT, Andrew

TITLE OF INVENTION: TANDEN FLUORESCENT PROTEIN CONSTRUCTS

FILE REFERENCE: REGEN1260-3

CURRENT APPLICATION NUMBER: US/10/057, 505

PRIOR FILING DATE: 2002-01-25

PRIOR FILING DATE: 1997-01-31

PRIOR APPLICATION NUMBER: US 09/396, 003

PRIOR FILING DATE: 1999-09-13

NUMBER OF SEQ ID NOS: 31

SOFTWARE: PatentIn version 3.1

SEQ ID NO 2

LENGTH: 238

TYPE: PRT

ORGANISM: Aequorea victoria

US-10-057-505-2

Query Match 100.0%; Score 1260; DB 13; Length 238;
Best Local Similarity 98.3%; Pred. No. 3e-112;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPLIVELDGVNGHKFSVSGEGDATTGKLTLLKFTCTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPLIVELDGVNGHKFSVSGEGDATTGKLTLLKFTCTTGKLPVWPPTL 60
QY 61 VTTXXYGVQCSRRYPDHMKRHDFFKSAMPEGVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
DB 61 VTTBSYGVQCSRRYPDHMKRHDFFKSAMPEGVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKIDPFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDGVSQQLAD 180
DB 121 NRIELKIDPFKEDGNILGHKLEYNNSHNVIYIMADKQNGIKVNFKIRHNIEDGVSQQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHNVLLFVYTAAGITTHGMDELYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHNVLLFVYTAAGITTHGMDELYK 238

RESULT 7

US-10-293-580-2
; Sequence 2, Application US/10293580
; Publication No. US20030170767A1
; GENERAL INFORMATION:

APPLICANT: Aurora Biosciences Corporation

APPLICANT: Cubitt, Andrew B.
TITLE OF INVENTION: Fluorescent Protein Sensors of Post-Translational Modifications
FILE REFERENCE: AURO1270 (08366/031001)
CURRENT APPLICATION NUMBER: US/10/293,580
CURRENT FILING DATE: 2002-11-12
PRIOR APPLICATION NUMBER: US/09/129,192
PRIOR FILING DATE: 1998-07-24
NUMBER OF SEQ ID NOS: 74
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 238
TYPE: PRT
ORGANISM: Aequorea
US-10-293-580-2

Query Match 100.0%; Score 1260; DB 14; Length 238;
Best Local Similarity 98.3%; Pred. No. 3e-112;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGEDATYKLTLEKTTGKLPVPMPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGEDATYKLTLEKTTGKLPVPMPTL 60
QY 61 VTTXXYGVQCFSRYPDHHKRDHFFKSAMPEGYQERTIFFKDDGNYKTRAIVKFEEDTLV 120
DB 61 VTTXSVGVQCFSRYPDHHKRDHFFKSAMPEGYQERTIFFKDDGNYKTRAIVKFEEDTLV 120
QY 121 NRLEKIDGIDFKEDGNILGHKLEYNNSHNYIMADKOKNGIKVNFKIRNIEDGVQVLAD 180
DB 121 NRLEKIDGIDFKEDGNILGHKLEYNNSHNYIMADKOKNGIKVNFKIRNIEDGVQVLAD 180
QY 181 HYQONTPIGDGPVLLPDNHVLTSTQSALSKDPNEKRDHMLVLEFVTAAGITGMDLEYK 238
DB 181 HYQONTPIGDGPVLLPDNHVLTSTQSALSKDPNEKRDHMLVLEFVTAAGITGMDLEYK 238

RESULT 8
US-10-293-580-74

Sequence 74, Application US/10293580
Publication No. US20030170767A1
GENERAL INFORMATION:
APPLICANT: Aurora Biosciences Corporation
APPLICANT: Cubitt, Andrew B.
TITLE OF INVENTION: Fluorescent Protein Sensors of Post-Translational Modifications
FILE REFERENCE: AURO1270 (08366/031001)
CURRENT APPLICATION NUMBER: US/10/293,580
CURRENT FILING DATE: 2002-11-12
PRIOR APPLICATION NUMBER: US/09/129,192
PRIOR FILING DATE: 1998-07-24
NUMBER OF SEQ ID NOS: 74
SOFTWARE: PatentIn version 3.0
SEQ ID NO 74
LENGTH: 238
TYPE: PRT
ORGANISM: Aequorea green fluorescent protein phosphorylation mutant
US-10-293-580-74

Query Match 100.0%; Score 1260; DB 14; Length 238;
Best Local Similarity 98.3%; Pred. No. 3e-112;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGEDATYKLTLEKTTGKLPVPMPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGEDATYKLTLEKTTGKLPVPMPTL 60
QY 61 VTTXXYGVQCFSRYPDHHKRDHFFKSAMPEGYQERTIFFKDDGNYKTRAIVKFEEDTLV 120
DB 61 VTTXSVGVQCFSRYPDHHKRDHFFKSAMPEGYQERTIFFKDDGNYKTRAIVKFEEDTLV 120
QY 121 NRLEKIDGIDFKEDGNILGHKLEYNNSHNYIMADKOKNGIKVNFKIRNIEDGVQVLAD 180
DB 121 NRLEKIDGIDFKEDGNILGHKLEYNNSHNYIMADKOKNGIKVNFKIRNIEDGVQVLAD 180

QY 181 HYQONTPIGDGPVLLPDNHVLTSTQSALSKDPNEKRDHMLVLEFVTAAGITGMDLEYK 238
DB 181 HYQONTPIGDGPVLLPDNHVLTSTQSALSKDPNEKRDHMLVLEFVTAAGITGMDLEYK 238

RESULT 9
US-10-457-982-2

Sequence 2, Application US/10457982
Publication No. US20030212265A1
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
APPLICANT: Miyawaki, Atsushi
APPLICANT: Llopis, Juan
APPLICANT: Machter, Rebekka M.
APPLICANT: Remington, S. James
TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
TITLE OF INVENTION: MEASURING THE PH OF A BIOLOGICAL SAMPLE
FILE REFERENCE: 07257/071001
CURRENT APPLICATION NUMBER: US/10/457,982
CURRENT FILING DATE: 2003-06-09
PRIOR APPLICATION NUMBER: US/09/602,641
PRIOR FILING DATE: 2000-06-22
PRIOR APPLICATION NUMBER: 09/172,063
PRIOR FILING DATE: 1998-10-13
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 238
TYPE: PRT
ORGANISM: Aequorea victoria
US-10-457-982-2

Query Match 100.0%; Score 1260; DB 15; Length 238;
Best Local Similarity 98.3%; Pred. No. 3e-112;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGEDATYKLTLEKTTGKLPVPMPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGEDATYKLTLEKTTGKLPVPMPTL 60
QY 61 VTTXXYGVQCFSRYPDHHKRDHFFKSAMPEGYQERTIFFKDDGNYKTRAIVKFEEDTLV 120
DB 61 VTTXSVGVQCFSRYPDHHKRDHFFKSAMPEGYQERTIFFKDDGNYKTRAIVKFEEDTLV 120
QY 121 NRLEKIDGIDFKEDGNILGHKLEYNNSHNYIMADKOKNGIKVNFKIRNIEDGVQVLAD 180
DB 121 NRLEKIDGIDFKEDGNILGHKLEYNNSHNYIMADKOKNGIKVNFKIRNIEDGVQVLAD 180
QY 181 HYQONTPIGDGPVLLPDNHVLTSTQSALSKDPNEKRDHMLVLEFVTAAGITGMDLEYK 238
DB 181 HYQONTPIGDGPVLLPDNHVLTSTQSALSKDPNEKRDHMLVLEFVTAAGITGMDLEYK 238

RESULT 10

US-10-724-178-2
Sequence 2, Application US/10724178
Publication No. US20040137528A1
GENERAL INFORMATION:
APPLICANT: Odyssey Thera, Inc.
APPLICANT: Michnick, Stephen
APPLICANT: Macdonald, Marlene
APPLICANT: Lamerdin, Jane
TITLE OF INVENTION: FRAGMENTS OF FLUORESCENT PROTEINS FOR PROTEIN-FRAGMENT
FILE REFERENCE: ODDY007
CURRENT APPLICATION NUMBER: US/10/724,178
CURRENT FILING DATE: 2003-12-01
PRIOR APPLICATION NUMBER: US 60/461,133
PRIOR FILING DATE: 2003-04-09
NUMBER OF SEQ ID NOS: 1067
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 238

TYPE: PRT
ORGANISM: Aequorea victoria
US-10-724-178-2

Query Match 100.0%; Score 1260; DB 16; Length 238;
Best Local Similarity 98.3%; Pred. No. 3e-112;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVNGHKFSVSGEGEDATYKLTLCFTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVNGHKFSVSGEGEDATYKLTLCFTTGKLPVWPPTL 60
QY 61 VTTXXYGVQCFSRYPDHKKRHDFFKSAMPEGVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
DB 61 VTTFSYGVQCFSRYPDHKKRHDFFKSAMPEGVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIIDFKEDGNILGHKLEYNNSHNVYIMADKQKNGIKVNFKIRHNIEDGVSQVLAD 180
DB 121 NRIELKGIIDFKEDGNILGHKLEYNNSHNVYIMADKQKNGIKVNFKIRHNIEDGVSQVLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSQDPNEKRDHVVLLXFTVAAGITTHGMDELYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSQDPNEKRDHVVLLXFTVAAGITTHGMDELYK 238

RESULT 11
US-10-757-624-2

Sequence 2, Application US/10757624
Publication No. US20040138420A1
GENERAL INFORMATION:
APPLICANT: Stubbs, Anne E.
APPLICANT: Jones, Anne E.
APPLICANT: Michael, Nigel P.
APPLICANT: Thomas, Nicholas
TITLE OF INVENTION: Fluorescent Proteins
FILE REFERENCE: PA0111
CURRENT APPLICATION NUMBER: US/10/757,624
CURRENT FILING DATE: 2004-01-14
PRIOR APPLICATION NUMBER: US 09/967,301
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: GB 0109858.1
PRIOR FILING DATE: 2001-04-23
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
LENGTH: 238
TYPE: PRT
ORGANISM: Aequorea victoria
US-10-757-624-2

Query Match 100.0%; Score 1260; DB 16; Length 238;
Best Local Similarity 98.3%; Pred. No. 3e-112;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVNGHKFSVSGEGEDATYKLTLCFTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVNGHKFSVSGEGEDATYKLTLCFTTGKLPVWPPTL 60
QY 61 VTTXXYGVQCFSRYPDHKKRHDFFKSAMPEGVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
DB 61 VTTFSYGVQCFSRYPDHKKRHDFFKSAMPEGVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIIDFKEDGNILGHKLEYNNSHNVYIMADKQKNGIKVNFKIRHNIEDGVSQVLAD 180
DB 121 NRIELKGIIDFKEDGNILGHKLEYNNSHNVYIMADKQKNGIKVNFKIRHNIEDGVSQVLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSQDPNEKRDHVVLLXFTVAAGITTHGMDELYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSQDPNEKRDHVVLLXFTVAAGITTHGMDELYK 238

RESULT 12
US-10-757-624-3

Sequence 3, Application US/10757624
Publication No. US20040138420A1
GENERAL INFORMATION:
APPLICANT: Stubbs, Anne E.
APPLICANT: Jones, Anne E.
APPLICANT: Michael, Nigel P.
APPLICANT: Thomas, Nicholas
TITLE OF INVENTION: Fluorescent Proteins
FILE REFERENCE: PA0111
CURRENT APPLICATION NUMBER: US/10/757,624
CURRENT FILING DATE: 2004-01-14
PRIOR APPLICATION NUMBER: US 09/967,301
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: GB 0109858.1
PRIOR FILING DATE: 2001-04-23
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn version 3.2
SEQ ID NO 3
LENGTH: 238
TYPE: PRT
ORGANISM: Artificial Sequence
OTHER INFORMATION: synthetic protein
US-10-757-624-3

Query Match 100.0%; Score 1260; DB 16; Length 238;
Best Local Similarity 98.3%; Pred. No. 3e-112;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVNGHKFSVSGEGEDATYKLTLCFTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVNGHKFSVSGEGEDATYKLTLCFTTGKLPVWPPTL 60
QY 61 VTTXXYGVQCFSRYPDHKKRHDFFKSAMPEGVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
DB 61 VTTFSYGVQCFSRYPDHKKRHDFFKSAMPEGVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIIDFKEDGNILGHKLEYNNSHNVYIMADKQKNGIKVNFKIRHNIEDGVSQVLAD 180
DB 121 NRIELKGIIDFKEDGNILGHKLEYNNSHNVYIMADKQKNGIKVNFKIRHNIEDGVSQVLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSQDPNEKRDHVVLLXFTVAAGITTHGMDELYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSQDPNEKRDHVVLLXFTVAAGITTHGMDELYK 238

RESULT 13
US-10-757-624-4

Sequence 4, Application US/10757624
Publication No. US20040138420A1
GENERAL INFORMATION:
APPLICANT: Stubbs, Anne E.
APPLICANT: Jones, Anne E.
APPLICANT: Michael, Nigel P.
APPLICANT: Thomas, Nicholas
TITLE OF INVENTION: Fluorescent Proteins
FILE REFERENCE: PA0111
CURRENT APPLICATION NUMBER: US/10/757,624
CURRENT FILING DATE: 2004-01-14
PRIOR APPLICATION NUMBER: US 09/967,301
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: GB 0109858.1
PRIOR FILING DATE: 2001-04-23
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn version 3.2
SEQ ID NO 4
LENGTH: 238
TYPE: PRT
ORGANISM: Artificial Sequence
OTHER INFORMATION: synthetic protein
US-10-757-624-4

Query Match 100.0%; Score 1260; DB 16; Length 238;
Best Local Similarity 98.3%; Pred. No. 3e-112;
Matches 234; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGELFTGVVPIVLELDGVDVNGHKFSVSGEGSDATYGLTLKFICTTGKLPVWPPTL 60
DB 1 MSKGELFTGVVPIVLELDGVDVNGHKFSVSGEGSDATYGLTLKFICTTGKLPVWPPTL 60
QY VTTXXYGVQCFSRYPDHMKRHDFFKSAPEGVQERTIFFKDDGNKTRAEVFEEDTLV 120
DB 61 VTTLTGVQCFSRYPDHMKRHDFFKSAPEGVQERTIFFKDDGNKTRAEVFEEDTLV 120
QY 121 NRTELKIDPFKEKGNILGHKLEYNNSHNVYIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
DB 121 NRTELKIDPFKEKGNILGHKLEYNNSHNVYIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLLEFVTAAGITTHGMDELYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLLEFVTAAGITTHGMDELYK 238

RESULT 14
US-09-920-922-4
; Sequence 4; Application US/09920922
; Patent No. US20020083488A1
; GENERAL INFORMATION:
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Sawano, Asako
; TITLE OF INVENTION: METHOD FOR MUTAGENESIS
; FILE REFERENCE: 11263-012001
; CURRENT APPLICATION NUMBER: US/09/920, 922
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: JP 2000-237166
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-09-920-922-4

Query Match 99.7%; Score 1256; DB 9; Length 238;
Best Local Similarity 97.9%; Pred. No. 7.1e-112;
Matches 233; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGELFTGVVPIVLELDGVDVNGHKFSVSGEGSDATYGLTLKFICTTGKLPVWPPTL 60
DB 1 MSKGELFTGVVPIVLELDGVDVNGHKFSVSGEGSDATYGLTLKFICTTGKLPVWPPTL 60
QY 61 VTTXXYGVQCFSRYPDHMKRHDFFKSAPEGVQERTIFFKDDGNKTRAEVFEEDTLV 120
DB 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAPEGVQERTIFFKDDGNKTRAEVFEEDTLV 120
QY 121 NRTELKIDPFKEKGNILGHKLEYNNSHNVYIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
DB 121 NRTELKIDPFKEKGNILGHKLEYNNSHNVYIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLLEFVTAAGITTHGMDELYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLLEFVTAAGITTHGMDELYK 238

RESULT 15
US-09-852-000-1
; Sequence 1; Application US/09852000
; Patent No. US20020099170A1
; GENERAL INFORMATION:
; APPLICANT: Osumi, Takashi
; APPLICANT: Tsukamoto, Toshiro
; APPLICANT: Tsukamoto, No. US20020099170A1
; APPLICANT: Yamasaki, Masatoshi
; TITLE OF INVENTION: GREEN FLUORESCENT PROTEINS AND BLUE FLUORESCENT

; TITLE OF INVENTION: PROTEINS
; FILE REFERENCE: 046124-5005-US
; CURRENT APPLICATION NUMBER: US/09/852, 000
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: JP 026418/1998
; PRIOR FILING DATE: 1998-01-23
; PRIOR APPLICATION NUMBER: US 09/121, 539
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: US 09/615, 655
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
; FEATURE:
; OTHER INFORMATION: Green fluorescent protein
US-09-852-000-1

Query Match 99.7%; Score 1256; DB 9; Length 238;
Best Local Similarity 97.9%; Pred. No. 7.1e-112;
Matches 233; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGELFTGVVPIVLELDGVDVNGHKFSVSGEGSDATYGLTLKFICTTGKLPVWPPTL 60
DB 1 MSKGELFTGVVPIVLELDGVDVNGHKFSVSGEGSDATYGLTLKFICTTGKLPVWPPTL 60
QY 61 VTTXXYGVQCFSRYPDHMKRHDFFKSAPEGVQERTIFFKDDGNKTRAEVFEEDTLV 120
DB 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAPEGVQERTIFFKDDGNKTRAEVFEEDTLV 120
QY 121 NRTELKIDPFKEKGNILGHKLEYNNSHNVYIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
DB 121 NRTELKIDPFKEKGNILGHKLEYNNSHNVYIMADKQNGIKVNFKIRHNIEDGSVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLLEFVTAAGITTHGMDELYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHVLLEFVTAAGITTHGMDELYK 238

Search completed: February 28, 2005, 14:42:17
Job time : 140 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 28, 2005, 14:21:18 ; Search time 41 Seconds
(without alignment)
558.527 Million cell updates/sec

Title: US-10-757-624-2-COPY
Perfect score: 1260
Sequence: 1 MSKGBELTGVVPLVELDG.....VLKRVTAAGTTHGMDLYK 238

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|---------------------|
| 1 | 1250 | 99.2 | 238 | 1 JQ1514 | green-fluorescent |
| 2 | 105 | 8.3 | 785 | 2 H72228 | hypothetical prote |
| 3 | 95.5 | 7.6 | 1224 | 1 ERHUAH | coatomer complex a |
| 4 | 93.5 | 7.4 | 887 | 2 E82590 | leucyl-tRNA synthet |
| 5 | 92.5 | 7.3 | 861 | 2 H64102 | leucine-tRNA ligase |
| 6 | 90.5 | 7.2 | 636 | 2 C79031 | hypothetical prote |
| 7 | 90.5 | 7.2 | 655 | 2 D83917 | DNA topoisomerase |
| 8 | 89.5 | 7.1 | 860 | 2 AC0582 | leucyl-tRNA synthet |
| 9 | 89 | 7.1 | 439 | 2 JH0414 | synaptogamin o-p65 |
| 10 | 88.5 | 7.0 | 797 | 2 JC4078 | protective surface |
| 11 | 88.5 | 7.0 | 808 | 2 F64102 | hypothetical prote |
| 12 | 88 | 7.0 | 281 | 2 AD2052 | hypothetical prote |
| 13 | 88 | 7.0 | 471 | 2 T27856 | hypothetical prote |
| 14 | 87.5 | 6.9 | 357 | 2 D71614 | hypothetical prote |
| 15 | 87 | 6.9 | 2573 | 2 G81355 | tRNA (tyracil-5-)-m |
| 16 | 87 | 6.9 | 370 | 2 E70390 | iron-sulfur cofact |
| 17 | 87 | 6.9 | 1152 | 2 AC1347 | probable peptidogl |
| 18 | 86.5 | 6.9 | 534 | 1 N1CLMA | nitrogenase (Fe 1, |
| 19 | 86.5 | 6.9 | 1151 | 2 AG1717 | probable peptidogl |
| 20 | 86 | 6.8 | 632 | 2 T06586 | DNA-binding protei |
| 21 | 86 | 6.8 | 1164 | 1 FCS6AG | IGA Fc receptor pr |
| 22 | 85.5 | 6.8 | 393 | 2 C64613 | conserved hypochet |
| 23 | 85.5 | 6.8 | 461 | 2 T06936 | photosystem II chl |
| 24 | 85.5 | 6.8 | 653 | 1 A11535 | fructose-bisphosph |
| 25 | 85.5 | 6.8 | 740 | 2 G95133 | neutraminidase, pro |
| 26 | 85.5 | 6.8 | 889 | 2 JC5576 | inter-alpha-trypsi |
| 27 | 85.5 | 6.8 | 1228 | 2 S60085 | nitrate reductase |
| 28 | 85.5 | 6.8 | 2222 | 1 A36028 | DNA-directed DNA p |
| 29 | 85 | 6.7 | 774 | 2 T39539 | alpha-amylase homo |

| | | | | | |
|----|------|-----|------|----------|---------------------|
| 30 | 84.5 | 6.7 | 322 | 2 T22410 | hypothetical prote |
| 31 | 84.5 | 6.7 | 1134 | 2 A60234 | IGA Fc receptor pr |
| 32 | 84 | 6.7 | 336 | 2 C64468 | hypothetical prote |
| 33 | 84 | 6.7 | 531 | 2 C95138 | hypothetical prote |
| 34 | 83.5 | 6.6 | 269 | 2 B35678 | enkephalin precurs |
| 35 | 83.5 | 6.6 | 310 | 2 S68225 | synergohymenotro |
| 36 | 83.5 | 6.6 | 564 | 2 B81317 | ABC-type transport |
| 37 | 83.5 | 6.6 | 583 | 2 T02209 | cellulose synthase |
| 38 | 83.5 | 6.6 | 1346 | 2 G71613 | hypothetical prote |
| 39 | 83 | 6.6 | 461 | 2 S41480 | photosystem II chl |
| 40 | 83 | 6.6 | 529 | 2 B86815 | hypothetical prote |
| 41 | 83 | 6.6 | 533 | 2 S47271 | AHNK-related proc |
| 42 | 83 | 6.6 | 578 | 1 I40794 | dihydroliipoamide d |
| 43 | 83 | 6.6 | 647 | 2 H89988 | hypothetical prote |
| 44 | 83 | 6.6 | 760 | 2 B84715 | probable oxygenase |
| 45 | 83 | 6.6 | 1092 | 2 S42798 | fibronectin-bindin |

ALIGNMENTS

RESULT 1
JQ1514
green-fluorescent protein [validated] - hydromedusa (Aequorea victoria)
C:Species: Aequorea victoria
C:Date: 03-Dec-1999 #sequence revision 03-Dec-1999 #text change 09-Jul-2004
C:Accession: J06092; JQ1514; P00335; S48693; S51330; S51331
R:Prasner, D.C.; Eckenrode, V.K.; Ward, W.W.; Prendergast, F.G.; Cornier, M.J.
Gene 111, 229-233, 1992
A:Title: Primary structure of the Aequorea victoria green-fluorescent protein.
A:Reference number: JQ1514; MUID:92175527; PMID:1347277
A:Accession: J06092
A:Molecule type: DNA
A:Residues: 1-107, 'S', 109-238 <PRA1>
A:Cross-References: UNIPROT:P42212; UNIPROT:Q17106; UNIPROT:Q17105; GB:M62654; NID:G15566
A:Accession: JQ1514
A:Molecule type: mRNA
A:Residues: 1-99, 'P', 101-140, 'L', 142-218, 'V', 220-238 <PRA2>
A:Cross-References: GB:M62653; NID:G155660; PIDN:AAA27721.1; PID:G155661
A:Accession: P00335
A:Molecule type: protein
A:Residues: 46-64; 74-122; 132-151; 154-183; 185-200 <PRA3>
R:Inouye, S.; Tsuji, F.I.
FEBS Lett. 351, 211-214, 1994
A:Title: Evidence for redox forms of the Aequorea green fluorescent protein.
A:Reference number: S48693; MUID:94364470; PMID:8082767
A:Accession: S48693
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-24, 'Q', 26-156, 'P', 158-171, 'K', 173-238 <INO>
A:Cross-References: GB:L29345; NID:9606383; PIDN:AAA58246.1; PID:9606384
R:Watkins, J.N.; Campbell, A.K.
submitted to the EMBL Data Library, January 1995
A:Reference number: S51330
A:Accession: S51330
A:Molecule type: mRNA
A:Residues: 1-13, 'V', 15-24, 'Q', 26-44, 'N', 46-153, 'G', 155-156, 'P', 158-171, 'K', 173-227, 'R', ;
A:Cross-References: EMBL:X83959; NID:9634008; PIDN:CAA8789.1; PID:9634009
A:Experimental source: clone gfp1
A:Accession: S51331
A:Molecule type: mRNA
A:Residues: 1-24, 'Q', 26-29, 'R', 31-83, 'L', 85-153, 'G', 155-156, 'P', 158-171, 'K', 173-208, 'Q', ;
A:Cross-References: EMBL:X83960; NID:9634010; PIDN:CAA8790.1; PID:9634011
A:Experimental source: clone gfp2
A:Yang, F.; Moss, L.G.; Phillips Jr., G.N.
submitted to the Brookhaven Protein Data Bank, August 1996
A:Reference number: A65692; PDB:1GFL
A:Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 'A', 2-79, 'R', 81-9;
A>Note: engineered sequence based on JQ1514, cloned and expressed in Escherichia coli
R:Yang, F.; Moss, L.G.; Phillips Jr., G.N.
Nat. Biotechnol. 14, 1246-1251, 1996
A:Title: The molecular structure of green fluorescent protein.
A:Reference number: A56953; MUID:98294543; PMID:9631087

A;Contents: annotation; X-ray crystallography, 1.9 angstroms
C;Comment: This protein is excited by the photoprotein aequorin (see PIR:AQJENV) emitting
C;Comment: The chromophore of this protein is formed by modification of Ser-dehydro-Tyr-
C;Genetics:
A;Gene: GFP
A;Intons: 69/3, 167/3
C;Superfamily: green-fluorescent protein
C;Keywords: chromoprotein; luminescence
F/65-67/Cross-link: 5-imidazolinone (Ser-Gly) #status experimental
F/66/Modified site: dehydrotyrosine (Tyr) #status experimental

Query Match 99.2%; Score 1250; DB 1; Length 238;
Best Local Similarity 96.6%; Pred. No. 3.8e-96;
Matches 230; Conservative 8; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVPLVLELDGVDNGHKFSVSGEGDATTYKLTLPKICTTGKLPVMPETL 60
DB 1 MSKGEELFTGVPLVLELDGVDNGHKFSVSGEGDATTYKLTLPKICTTGKLPVMPETL 60
QY 61 VTTXXYVQCFSRYPDHMKHDFPKSAMPEGVQERTIFPKDDGNVYTRAEVKEGDTLV 120
DB 61 VTTFSYVQCFSRYPDHMKQHDFPKSAMPEGVQERTIFPKDDGNVYTRAEVKEGDTLV 120
QY 121 NRLEKIDPKKEGNILGHLENYNSHNYIMADKQNGIKVNFKIRHNIEDGXVQLAD 180
DB 121 NRLEKIDPKKEGNILGHLENYNSHNYIMADKQNGIKVNFKIRHNIEDGXVQLAD 180
QY 181 HYQONTPIGDPVLLPDNHYLSTQSALSKDPNEKRDHMLLVFTYAGITGHMDELYK 238
DB 181 HYQONTPIGDPVLLPDNHYLSTQSALSKDPNEKRDHMLLVFTYAGITGHMDELYK 238

RESULT 2

H72228
hypothetical protein TM1624 - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
R/Nelson, K.E.; Clayton, R.A.; Gail, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; WUID:99287316; PMID:10360571
A;Accession: H72228
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-785 <ARN>
A;Cross-references: UNIPROT:Q9X1Y9; GB:AE001806; GB:AE000512; NID:94982196; PIDN:AAD3669
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1624

Query Match 8.3%; Score 105; DB 2; Length 785;
Best Local Similarity 19.3%; Pred. No. 1.2;
Matches 45; Conservative 35; Mismatches 69; Indels 84; Gaps 7;

QY 2 SKGEELFTGVPLVLELDGVDNGHKFSVSGEGDATTYKLTLPKICTTGKLPVMPETLV 61
DB 15 NBSGRFSEGVVPGVQAD-----LVKGLLPHVYGM- 46
QY 62 TTXXYVQCFSRYPDHMKHDFPKSAMPEGVQERTIFPKDDGNVYTRAEVKEGDTLV 121
DB 47 -----NEDLFKEIEDREWIYERFEFEKEDVKEGERDVLVEGVDTLIS 88
QY 122 RIELKIDPKKEGNILGHLENYNSHNYIMADKQNGIKVNFKIRHNIEDGXVQLADH 181
DB 89 DYLKNGVYL--GSTEDMFLERAFDVTNVL-----KEKNHLKYIK-----SPIRVPKI 134
QY 182 YOONTPIGDPVLLPDNHYLSTQSALSKDPNEKRDHMLLVFTYAGITGHMDELYK 234
DB 135 LKQNYVGLGP-----EDP-----IKGIRKQYISGM 163

RESULT 3

ERHUAH
coatomer complex alpha chain homolog - human
N/Alternate names: HEP-COP; xenopsin homolog; xenopsin-related peptide precursor
N/Contains: xenin 25; xenopsin-related peptide
C;Species: Homo sapiens (man)
C;Date: 10-May-1996 #sequence_revision 08-Nov-1996 #text_change 09-Jul-2004
R;Chow, V.T.K.; Quek, H.H.
Gene 169, 223-227, 1996

A;Title: HEP-COP, a novel human gene whose product is highly homologous to the alpha-sub
A;Reference number: J04668; WUID:96194806; PMID:8647451
A;Accession: J04668
A;Molecule type: mRNA
A;Residues: 1-1224 <CHO>
A;Cross-references: UNIPROT:P53621; GB:U24105; NID:g1638873; PIDN:AAB70879.1; PID:g10023
A;Experimental source: Hep3B hepatocellular carcinoma cell
R;Feurle, G.E.; Hamacher, G.; Kusilek, R.; Meyer, H.E.; Metzger, J.W.
J. Biol. Chem. 267, 22305-22309, 1992
A;Title: Identification of xenin, a xenopsin-related peptide, in the human gastric mucosa
A;Reference number: A44317; WUID:93054515; PMID:1429581

A;Accession: A44317
A;Molecule type: protein
A;Residues: 1-25 <FEU>
A;Experimental source: gastric mucosa
A;Note: sequence extracted from NCBI backbone (NCBIP:117018)
A;Note: plasma levels of xenin 25, as determined by immunoassay, rise after meals and it
C;Genetics:
A;Gene: GDB:COPA; HEP-COP
A;Cross-references: GDB:4642787; OMIM:601924
A;Map position: 1q23-1q25
C;Superfamily: Yeast coatomer complex alpha chain: WD repeat homology
C;Keywords: duplication; hormone; plasma; stomach
F/1-25/Product: xenin 25 #status experimental <XNP>
F/38/Domain: WD repeat homology <WD1>
F/17-25/Product: xenopsin-related peptide #status predicted <XNP>
F/47-80/Domain: WD repeat homology <WD2>
F/89-122/Domain: WD repeat homology <WD3>
F/131-164/Domain: WD repeat homology <WD4>
F/201-234/Domain: WD repeat homology <WD5>
F/245-278/Domain: WD repeat homology <WD6>

Query Match 7.6%; Score 95.5; DB 1; Length 1224;
Best Local Similarity 23.0%; Pred. No. 14;
Matches 56; Conservative 36; Mismatches 77; Indels 75; Gaps 13;

QY 44 LKPICTTG-----KLVPWPVLVTTXXYG--VQCFSR----- 73
DB 549 IKYAVTYGDHGIIRTLDP---LYTVKGNVYCLDRECRPVLITPTEFKFLALIN 605
QY 74 --YPD--HKRH-----DFKSAWPEGVQERTIFPKDDGNVYTRAEVKEGDTLVNR 123
DB 606 RKYDEVLHVNRAKLVGSGITAVYLOKKGPEVALHFVDE--KTRFSLALECGNITEL 662
QY 124 E-LKSIDPKKEGNILGHLENYNSHNYIMADKQNGIKVNF-----KIRHNIED 173
DB 663 EAAKALDQKCEKKEGEVALVQGNHGIYEMCGQRKRNKNDKVSFLVLTGNLEKAKMKKI 722
QY 174 GXVQ--LADHYQONTPIGDPVLLPDNHYLSTQSALSKDPNEKRDHMLLVFTYAGIT 231
DB 723 AEIRKDMGSHYONALYLD-----VSEBRVILKNGCGKS-----LAVLTAA--TH 765
QY 232 GWDE 235
DB 766 GLDE 769

RESULT 4

E82590
leucyl-tRNA synthetase XP2176 [imported] - Xylella fastidiosa (strain 945c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000

C;Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 7.3%; Score 92.5; DB 2; Length 861;
Best Local Similarity 24.3%; Pred.No.15;
Matches 45; Conservative 28; Mismatches 65; Indels 47; Gaps 9;

Dy TTGKLPVMPPLVTXXXYGVCCFSRPYDHMKRHPFKSAMPEGVOERTIFFKDDGNVKT 108
||| :||| :
Db TGDKLPI-WANFVLNHTGTGAWAVPAHDQRD--FEFAOKYSLLPQIVAPLADDEIDL 370
314 ||| :||| :
Oy RAAYKEGGDTLVNRILELKIDPEKDGNILGHKLKYNNSHNYIMADK-QKNGI---KVN 164
109 ||| :||| :
Db TKQAARVEHGKLVNSDFDGKNF--DCAFNFG-----IADKLEKLGVCGRQVN 414
371 ||| :||| :
Oy FKIRH-----NIEDGXVOLADHYOQNTPIGDGPVLLPPNHYL-STOSAL 207
165 ||| :||| :
Db YRLRWGVSRRORYWGAPAFMPLTLENGDVVA-----PMEDLPILPEDVNVMDGVKS 467
415 ||| :||| :
Oy SKDPN 212
208 ||| :||| :
Db NADPN 472
468 ||| :||| :

RESULT 6
C70031
hypothetical protein yycB - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: C70031
R;Kunst, F.; Ogasawara, S.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteiro,
A.C.; Bron, S.; Brouillet, N.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi,
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Filtz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallero,
leech, J.; Harwood, C.R.; Hensut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hulio, M.F.;
Koester, P.; Koningsstern, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel,
Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Poterelle,
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A.; Authors: Schlecht, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror,
keuchl, M.; Tanakoshi, A.; Tanaka, T.; Terpestra, P.; Togomi, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A;Authors: Yoshikawa, H.F.; Zumestein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A65580; MUID:98044033; PMID:9384377
A;Accession: C70031
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-636 <KUN>
A;Cross-references: UNIPROT:O06966; GB:Z99121; GB:AL009126; NID:g2635827; PIDN:CABI5489.1
A;Experimental source: strain 168
C;Genetics:
A;Gene: yycB
C;Superfamily: Bacillus subtilis hypothetical protein yycB

Query Match 7.2%; Score 90.5; DB 2; Length 636;
Best Local Similarity 22.2%; Pred.No.15;
Matches 58; Conservative 39; Mismatches 99; Indels 65; Gaps 14;

Oy BEL--FTGVVPLVLVDGYNHGKRSVSGEGSDATYYKLTLCFTCTGKLPVWPPLVLT 62
||| :||| :
Db BELPFLPFLHLTLRNNSGFVD---QLANGVIDETOKVKIKEL-----TNIN 401
356 ||| :||| :
Oy TXXYGVCCSRYP---DHMKRHDFPKSAMPEG---YVOERTIFFKDDGNVKT 114
63 ||| :||| :
Db ASMRGIE--SRKEHPIDHLKHGQFWALXDVGEIYDVYKSLISLKLDTDETAKLIIV 459
402 ||| :||| :
Oy EGDTLVNRILEKGIDPEKDGNILGHKLKYNNSHNYIMADKONG--IKVN----- 164
115 ||| :||| :
Db DGHSIVQML-----NALSKKKGSFYKSDIY-FPGSGSGERTIOVNLSSAVRIY 507
460 ||| :||| :
Oy EKIRHNIEGQXVOLADHYOQNTPIGDGPVLLPPNHYSTOSAL---SKDPNEKDDHM 218
||| :||| :
||| :||| :

Db 508 QNGKIVEDMEDAISKYKVSORH-----DEDFIDKQAITTAIHMEENPSHY 557
Qy 219 VL-LKFTVAGITGHMDLYK 238
Db 558 AFDLQFRLAAQFSTHTPDLK 578

RESULT 7

DNA topoisomerase IV subunit B BH2140 [imported] - Bacillus halodurans (strain C-125)
D83917
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C/Accession: D83917
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: D83917
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-655 <STO>
A/Cross-references: UNIPROT:Q9KAZ6; GB:A001514; GB:BA000004; NID:g10174613; PIDN:BA058
A/Experimental source: strain C-125
C/Genetics:
A/Gene: BH2140
C/Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain B

Query Match 7.2%; Score 90.5; DB 2; Length 655;

Best Local Similarity 21.3%; Pred. No. 15; Mismatches 83; Indels 81; Gaps 11;

Matches 54; Conservative 36; Mismatches 83; Indels 81; Gaps 11;

Qy 21 DVNGHK--FSVSGEGGDATYKLTGKFLCTTG--KLPPWPPLVTTXXYGVQCFER-- 73

Db 63 NVTIHKQSVSVNDEGRG-----MPTGMHLKGPTEPVILTLVHAGSKRGG 110

Qy 74 YPDHMKRHDPFKSM-----PEGYOER-----TI 98

Db 111 YATSGLHGVAGVAVNALSEMLVEIKRDWVYEQREFNGSKSTLEKKKTRQCTTT 170

Qy 99 FFKDDG-----NYKTRAEVKEGDTLVNRIELKIDPEKDNILGHLEVNYSNHY 150

Db 171 HFRPDPVSTNNVTTLSERLERAAFLKGLKIELVDRDPTKEVFF--YEDGKAEVE 229

Qy 151 YIMADK-----QKNGIKVNFKIRNIEDGVQVLADHYQONTPIGDGVLTPDNH 199

Db 230 YLNEDEKTLHPVVFENESNGIEIEFAFQFN--DGYTEVTLSPVNVVTRTDG-----GTH 282

Qy 200 YLSTQSLSKDPPE 213

Db 283 ELGAKTAMTRAVNE 296

RESULT 8

AC0582
leucyl-tRNA synthetase [imported] - Salmonella enterica subsp. enterica serovar Typhi (S
C/Species: Salmonella enterica subsp. enterica serovar Typhi
A/Note: this species has also been called Salmonella typhi
C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C/Accession: AC0582
R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.W.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Garra, P.
Nature 413, 848-852, 2001
A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A/Reference number: AB0502; MUID:21534947; PMID:11677608
A/Accession: AC0582
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-860 <PAR>
A/Cross-references: GB:AL513382; PIDN:CAD05125.1; PID:g16501899; GSPDB:GN00176
C/Genetics:
A/Gene: SRY0699

C/Superfamily: leucine-tRNA ligase

Query Match 7.1%; Score 89.5; DB 2; Length 860;

Best Local Similarity 22.8%; Pred. No. 27; Mismatches 41; Conservative 23; Mismatches 77; Indels 39; Gaps 6;

Qy 49 TTGKLPVMPPLVTTXXYGVQCFERYPDMMKRHDFFPKSAMPEGVOERTIFFDDGNVKT 108

Db 314 TGEIIPV-MAANVLMWEYGVGAVMAVPGHDQRD--YEFASKYGLTTPVTLADSGSEPL 370

Qy 109 RAEVKEGDTLVNRIELKIDPEKDNILGHLEVNYSNHYIMADKQNGIKVNFKIR 168

Db 371 SEQALTEKGVLFNSGFBFDGLAFPAFNAIADKL-----AEKGVGERKVVYRLR 418

Qy 169 H-----NDEGVQVLADHYQONTPIGDGVLTPDNHYL-STQSLSKSP 211

Db 419 DWGVSQRQRYWADIPWVTLDEGT-----LTPEDQLPVILFEDVVMGDTISPTKADP 471

RESULT 9

JH0414
synaptotagmin o-p65-B - electric ray (Discothyrea ommata)
N/Alternate names: synaptic vesicle protein o-p65-B
C/Species: Discothyrea ommata
C/Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 09-Jul-2004
C/Accession: JH0414; P80223
R/Wendland, B.; Miller, K.G.; Schilling, J.; Scheller, R.H.
Neuron 6, 993-1007, 1991
A/Title: Differential expression of the p65 gene family.
A/Reference number: JH0413; MUID:91273991; PMID:2054189
A/Accession: JH0414
A/Molecule type: mRNA
A/Residues: 1-439 <MEN>
A/Cross-references: UNIPROT:P24506; GB:M64276; NID:g213110; PIDN:AAA49228.1; PID:g213111
A/Experimental source: electric organ
A/Accession: P80223
A/Molecule type: protein
A/Residues: 'MLV', 26-34; 'XX', 194-199; 'X', 201-206; 'X', 322-332; 'D', 334-337 <MEN1>
C/Superfamily: synaptotagmin; protein kinase C C2 region homology
C/Keywords: glycoprotein; membrane protein; synaptic vesicle
F/75-101/Domain: hydrotrophic <HD>
F/153-266/Domain: protein kinase C C2 region homology <KC2A>
F/284-399/Domain: protein kinase C C2 region homology <KC2B>
F/6,46/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 7.1%; Score 89; DB 2; Length 439;

Best Local Similarity 19.2%; Pred. No. 12; Mismatches 55; Conservative 48; Mismatches 92; Indels 92; Gaps 12;

Qy 15 LVELDGVNGHKFSVSGEGGDATYKLTGKFLPV-FWPTLVTTXXYGV----- 68

Db 38 MNPIDTGDNSTEAGVPEBKND-VFEKLEKEMNELQKTLPRPMALIAIATYSGILLTLC 96

Qy 69 -----QCFSRYPDMMKRHDFFPKSAMPEGVOERTIFFDDGNVKTAEVKEG----- 116

Db 97 CLCTCKKCCCKKKKKKKKKKKGK-----KNDINMK---DVKSGGQDDDD 138

Qy 117 -DTLVNRIELKIDPEKDNILGHLEVNYSNHY----- 148

Db 139 DAEVTEGTEDEBEKEKEKEKLGKTFQSLDVPQANQTLVGIIOAAELPALDMGTSDDY 198

Qy 149 -NVIYIMADKQNGIKVNF-----FKIRNIEDGVQVLADHYQONTPI 188

Db 199 VAVFPLPDKKKYERKVKQKTLNPTFNESFVFKVYQELGKTLMMNAVYDFPRFKHDCI 258

Qy 189 GGPVLLPD-----NHYLSTQSLSKDPPEKRDHMLVLAFTVTAAG 228

Db 259 GQVTVLMTKVVDIGQQLLEWRDLSEAKKEPEKLGICISLRVVPFAG 305

RESULT 10

JC4078
protective surface antigen D-15 precursor - Haemophilus influenzae (type b)

C:Species: Haemophilus influenzae
A:Vary: type b
C:Date: 30-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 09-Jul-2004
C:Accession: J04078
R:Flack, P.S.; Loosmore, S.; Chong, P.; Thomas, W.R.
Gene 156, 97-99, 1995
A:Title: The sequencing of the 80-kDa D15 protective surface antigen of Haemophilus infl
A:Reference number: J04078; MUID:95255676; PMID:7737523
A:Accession: J04078
A:Molecule type: DNA
A:Residues: 1-797 <FLA>
A:Cross-references: UNIPROT:P46024; GB:U13961; NID:9537447; PIDN:AAA85645.1; PID:9537448
A:Experimental source: type b
C:Superfamily: protective surface antigen D-15
C:Keywords: surface antigen
P.1-19/Domains: signal sequence #status predicted <SIG>
P.20-797/Product: protective surface antigen D-15 #status predicted <MAT>

Query Match 7.0%; Score 88.5; DB 2; Length 797;
Best Local Similarity 21.5%; Pred. No. 29;
Matches 47; Conservative 35; Mismatches 74; Indels 63; Gaps 11;

Qy 64 XYYGVQCFRRYPDHMKRHP-----FKSAMPEGYQOE-----RTI 98
Db 427 IGYTESGISYQASVQDNFLGTGAAVSIAGTKNDYGTSVNLGYTEPYTKDGVSLGAV 486
Qy 99 FKPDGNYTRAIVKFEQGTLYNRILEKIDKEKDNIT---LGH-----KLEYNVNS 147
Db 487 FFENYDNSKSDTSSNYKRTTYSNVTL--GFPVNNNSYYVGLGHTYKNSINPALRYN--- 542
Qy 148 HHVYIMADKQK-NGIKVNFKIRHNIEDGAVQLADHYHQ-----NTPIGDGPVLL 195
Db 543 RLVIYQSMKFKNGIKITN-----DPPSFGMNYNSLNRGYPFKYKASLG-GRVTI 593
Qy 196 P--DNHYLSTQSALSKDPNEKRDHVLXFTVAAGITGG 232
Db 594 PGSDNKYYKLADVOGFYPLDRDHLVWVSAKASAGYANG 632

RESULT 11
F64102
protective surface antigen D-15 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 18-Sep-1998
C:Accession: F64102
R:Feleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
; Gocayne, J.D.; Scott, J.; Shinkley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghegan, N.S.M.
Science 269, 496-512, 1995
A:Author: Gnehm, C.L.; McDonald, L.A.; Smill, K.V.; Frazer, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: F64102
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-808 <TIGR>
A:Cross-references: GB:U42023; TIGR:HI0917
C:Superfamily: protective surface antigen D-15
C:Keywords: surface antigen

Query Match 7.0%; Score 88.5; DB 2; Length 808;
Best Local Similarity 21.5%; Pred. No. 30;
Matches 47; Conservative 35; Mismatches 74; Indels 63; Gaps 11;

Qy 64 XYYGVQCFRRYPDHMKRHP-----FKSAMPEGYQOE-----RTI 98
Db 440 IGYTESGISYQASVQDNFLGTGAAVSIAGTKNDYGTSVNLGYTEPYTKDGVSLGAV 499
Qy 99 FKPDGNYTRAIVKFEQGTLYNRILEKIDKEKDNIT---LGH-----KLEYNVNS 147
Db 500 FFENYDNSKSDTSSNYKRTTYSNVTL--GFPVNNNSYYVGLGHTYKNSINPALRYN--- 555
Qy 148 HHVYIMADKQK-NGIKVNFKIRHNIEDGAVQLADHYHQ-----NTPIGDGPVLL 195

Db 556 RLVIYQSMKFKNGIKITN-----DPPSFGMNYNSLNRGYPFKYKASLG-GRVTI 606
Qy 196 P--DNHYLSTQSALSKDPNEKRDHVLXFTVAAGITGG 232
Db 607 PGSDNKYYKLADVOGFYPLDRDHLVWVSAKASAGYANG 645

RESULT 12
AD2052
hypothetical protein all1970 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AD2052
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Saemoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anat
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AD2052
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-281 <KUR>
A:Cross-references: UNIPROT:O8YK5; GB:BA000019; PIDN:BAW73669.1; PID:gl17131060; GSPDB:G
A:Experimental source: strain PCC 7120
A:Genetics: all1970

Query Match 7.0%; Score 88; DB 2; Length 281;
Best Local Similarity 24.8%; Pred. No. 7.9;
Matches 31; Conservative 26; Mismatches 46; Indels 22; Gaps 6;

Qy 106 YKTRAEVKEEGDTLYNRILEKGI---DPKEDGNILGHLEYNVNSHHVYIMADKQKGI 161
Db 46 FKARS-LOSSSDILNLPKYNQLQNRPDWKDEYIVAAEGYKGE-----SSKRIKSI 99
Qy 162 KYNFKI--RHNIEDGAVQLADH-----YQONTPIGQ---FVLLPNNHYLSTQSALSK 209
Db 100 KFGSLQERGDIVNQLDLNVLIPDEPTQTDTPPSGQNTPLMRDVKFISQPKTITK 159

Qy 210 DPNEK 214
Db 160 QQKER 164

RESULT 13
T27856
hypothetical protein ZK418.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T27856
R:Fulton, L.
submitted to the EMBL Data Library, April 1994
A:Description: The sequence of C. elegans coamid ZK418.
A:Reference number: Z20430
A:Accession: T27856
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-471 <FUL>
A:Cross-references: UNIPROT:Q24488; EMBL:U00047; PIDN:AAA50686.1; CESP:ZK418.2
A:Experimental source: strain Bristol N2
A:Genetics: ZK418.2
A:Interons: 41/3; 59/1; 139/1; 214/3; 241/3; 295/3; 329/3; 381/3; 454/2
C:Superfamily: Caenorhabditis elegans hypothetical protein ZK418.2

Query Match 7.0%; Score 88; DB 2; Length 471;
Best Local Similarity 22.9%; Pred. No. 16;
Matches 48; Conservative 36; Mismatches 80; Indels 46; Gaps 10;

Qy 46 FICTTGKLPVWPPTLVTTXXYGVQCFRRYPDHMKRHP-----FKSAMPE-----GYQERTI 98
Db 160 QQKER 164

Db 140 FLCKFSVLPV-----SHY-----RPIEMKGGVPTFKREVELNDVKTNNEM 186
QY 99 FF---KDD-GNYKTRAEVKFE-----GDTLVNRIELKGDPEKEDGNILGHKLEYNN 146
Db 187 FFDLLDNDGNSGCDTLVLTDSKRTMTKGVALLIIVGMTTGKEVFEIAGNYVNLMLKSAHV 246
QY 147 SHNVYIMADKQKNGIKV-NFKIRHNIEDGXVOLADHYQONTPIGDGPVLL----- 195
Db 247 VPSNYKLDKLFTEGEKVPQCDWVFVYKDGSP--SDFVHVPVLADSSIVRAAVVSHMSI 304
QY 196 -PDNHYISTOSALSKDPNEKRDHMYLLXFV 224
Db 305 GPQIAMISHENRLIFNTLKSRDWHEVILTYI 334

RESULT 14

DJ1614
hypothetical protein PFB0460c - malaria parasite (Plasmodium falciparum)
C/Species: Plasmodium falciparum
C/Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004
C/Accession: D71614
R/Gardner, M.J.; Tertelijn, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
Pereira, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1128-1132, 1998
A/Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A/Reference number: A71600; MUID:99021743; PMID:9804551
A/Accession: D71614
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-2573 <GAR>
A/Cross-references: UNIPROT:O96185; GB:AE001396; GB:AE001362; MID:G3845188; PIDN:AAC7188
A/Experimental source: clone 3D7
C/Genetics:
A/Gene: PFB0460c

Query Match 6.9%; Score 87.5; DB 2; Length 2573;
Best Local Similarity 26.2%; Pred. No. 1.7e+02;
Matches 34; Conservative 30; Mismatches 53; Indels 13; Gaps 5;

QY 93 VOERTIEFKD--DGNVKTAEVKEEGPTLVNRIELKGDPEKEDGNILGHKLEYN--YNSH 148
Db 126 LKKEITLCKDIKGSNDPMDEISLFDKDWVDKELK--DPEKSLKIKKKEVYNFTYNNK 183
QY 149 NVYIMADKQKNGIKVNFKIRHNIEDGXVOLADHYQONTPIGDGPVLLPDNHYISTOSALS 208
Db 184 NLAHKENKKDKKKKIKHNNDENNMM---IYYKNI---DKHYIADNNVVHILNDIN 236
QY 209 KDPNEKRDHM 218
Db 237 TYLKRERDYM 246

RESULT 15

G81355
tRNA (uracil-5-)-methyltransferase (EC 2.1.1.35) Cj0831c [imported] - Campylobacter jejuni
C/Species: Campylobacter jejuni
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C/Accession: G81355
R/Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whittread, S.; Barrer
Nature 403, 665-668, 2000
A/Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hfg
A/Reference number: A81250; MUID:20150912; PMID:10688204
A/Accession: G81355
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-357 <PAR>
A/Cross-references: UNIPROT:Q9PP92; GB:AL139076; GB:AL111168; MID:G6968128; PIDN:CAB7309
A/Experimental source: serotype O2, strain NCTC 11168
C/Genetics:
A/Gene: tRNA; Cj0831c
C/Keywords: methyltransferase, S-adenosylmethionine

Query Match 6.9%; Score 87; DB 2; Length 357;
Best Local Similarity 24.8%; Pred. No. 13;
Matches 30; Conservative 18; Mismatches 39; Indels 34; Gaps 5;

QY 79 KRHDFKSMPEGYVOERTIFFKDDGNVKTAEVKF--EGDTLV-----NRIELKG 127
Db 14 EKHSFKKTYFKERYTDFKLFAKDKGVRTRAELSFYHENDTLFVAMPDPKSKKKYIIIEY 73
QY 128 IDPEKED-----GNILGHKLEYNNSHNVYIMADKQKNGIKVNFKIRHNIE 172
Db 74 LDFADEKICAFMRPLLLEYLRQDKLKEKL-----FGVEFLTTCOE--ISTILLHKNIE 125
QY 173 D 173
Db 126 D 126

Search completed: February 28, 2005, 14:30:21
Job time : 45 secs

| | | | | | | | |
|----|-------|-------|-----|---|-----------|--------|--------------|
| 1 | 1260 | 100.0 | 238 | 2 | Q8GHE2 | Q8GHE2 | azotobacter |
| 2 | 1256 | 99.7 | 238 | 1 | GFP_AEQYT | P42312 | aequorea vi |
| 3 | 1256 | 99.7 | 238 | 2 | Q71R9Y | Q71R9Y | azotobacter |
| 4 | 1255 | 99.6 | 238 | 2 | Q8GHE4 | Q8GHE4 | azomonas ag |
| 5 | 1254 | 99.5 | 238 | 2 | Q8GHE3 | Q8GHE3 | azotobacter |
| 6 | 1253 | 99.4 | 238 | 2 | Q9J325 | Q9J325 | aequorea vi |
| 7 | 1214 | 96.3 | 238 | 2 | Q17105 | Q17105 | aequorea vi |
| 8 | 1199 | 95.2 | 238 | 2 | Q17106 | Q17106 | aequorea vi |
| 9 | 1186 | 94.1 | 238 | 2 | Q6Y6Z0 | Q6Y6Z0 | aequorea co |
| 10 | 1097 | 87.1 | 238 | 2 | Q8WP95 | Q8WP95 | aequorea ma |
| 11 | 1097 | 87.1 | 238 | 2 | Q8WTC6 | Q8WTC6 | aequorea ma |
| 12 | 1093 | 86.7 | 238 | 2 | Q8WTC4 | Q8WTC4 | aequorea ma |
| 13 | 1091 | 86.6 | 238 | 2 | Q8WTD0 | Q8WTD0 | aequorea ma |
| 14 | 1090 | 86.5 | 238 | 2 | Q8WTC8 | Q8WTC8 | aequorea ma |
| 15 | 1090 | 86.5 | 238 | 2 | Q8WTC9 | Q8WTC9 | aequorea ma |
| 16 | 1086 | 86.2 | 238 | 2 | Q8WTC5 | Q8WTC5 | aequorea ma |
| 17 | 1084 | 86.0 | 238 | 2 | Q8WTC7 | Q8WTC7 | aequorea ma |
| 18 | 630 | 56.0 | 234 | 2 | Q6RY67 | Q6RY67 | phialidium |
| 19 | 465 | 36.9 | 225 | 2 | Q6RY65 | Q6RY65 | anthomedusa |
| 20 | 433 | 34.4 | 262 | 2 | Q6RY66 | Q6RY66 | anthomedusa |
| 21 | 256.5 | 20.4 | 225 | 2 | Q9SUD7 | Q9SUD7 | montastraea |
| 22 | 256.5 | 20.4 | 225 | 2 | Q7Z0M5 | Q7Z0M5 | montastraea |
| 23 | 250 | 19.8 | 225 | 2 | Q96J35 | Q96J35 | montastraea |
| 24 | 248 | 19.7 | 230 | 2 | Q66PW1 | Q66PW1 | scoliymia cu |
| 25 | 247.5 | 19.6 | 226 | 2 | Q8T6V0 | Q8T6V0 | dendronephr |
| 26 | 246.5 | 19.6 | 225 | 2 | Q7Z0V9 | Q7Z0V9 | montastraea |
| 27 | 245 | 19.4 | 232 | 2 | Q6RY64 | Q6RY64 | anthomedusa |
| 28 | 244 | 19.4 | 225 | 2 | Q8I6J8 | Q8I6J8 | trachyphyll |
| 29 | 241.5 | 19.2 | 226 | 2 | Q9UC73 | Q9UC73 | clavularia |
| 30 | 240.5 | 19.1 | 225 | 2 | Q6USK3 | Q6USK3 | montastraea |
| 31 | 236 | 18.7 | 225 | 2 | Q66ND3 | Q66ND3 | montastraea |

| | | | | | | |
|----|-------|------|-----|---|--------|--------------------|
| 32 | 236 | 18.7 | 225 | 2 | Q720W4 | Q720W4 montastraea |
| 33 | 235.5 | 18.7 | 225 | 2 | Q6R5F1 | Q6R5F1 montastraea |
| 34 | 223.5 | 17.7 | 223 | 2 | Q6R5F5 | Q6R5F5 astrangia l |
| 35 | 220 | 17.5 | 224 | 2 | Q6MU48 | Q6MU48 montastraea |
| 36 | 220 | 17.5 | 227 | 2 | Q66ND2 | Q66ND2 montastraea |
| 37 | 220 | 17.5 | 227 | 2 | Q66ND4 | Q66ND4 montastraea |
| 38 | 220 | 17.5 | 227 | 2 | Q66ND5 | Q66ND5 montastraea |
| 39 | 219.5 | 17.4 | 221 | 2 | Q66PV1 | Q66PV1 acropora te |
| 40 | 219 | 17.4 | 221 | 2 | Q66PV5 | Q66PV5 acropora m |
| 41 | 218.5 | 17.3 | 221 | 2 | Q66PV8 | Q66PV8 acropora ac |
| 42 | 217 | 17.2 | 227 | 2 | Q720W6 | Q720W6 montastraea |
| 43 | 216.5 | 17.2 | 234 | 2 | Q720W7 | Q720W7 montastraea |
| 44 | 216.5 | 17.2 | 259 | 2 | Q8WMA2 | Q8WMA2 agaricia fr |
| 45 | 214 | 17.0 | 229 | 2 | Q9U6Y6 | Q9U6Y6 anemonia ma |

ALIGNMENTS

| ID | Q8GHE2 | PRELIMINARY; | PRT; | 238 AA. |
|--------|---|---|------|---------|
| Q8GHE2 | Q8GHE2 | | | |
| DT | 01-MAR-2003 | (TrEMBLrel. 23, Created) | | |
| DT | 01-MAR-2003 | (TrEMBLrel. 23, Last sequence update) | | |
| DT | 01-OCT-2003 | (TrEMBLrel. 25, Last annotation update) | | |
| DE | Green fluorescence protein. | | | |
| GN | Name=2289Gfp; | | | |
| OS | Azotobacter vinelandii. | | | |
| OC | Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; | | | |
| OC | Pseudomonadaceae; Azotobacter. | | | |
| OX | NCBI_TaxID=354; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=DSM2289; | | | |
| RA | Koranyi P, Berenyi M, Burg K.; | | | |
| RL | Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases. | | | |
| DR | EMBL; AF344408; AAAB6140.1; - | | | |
| DR | HSSP; P42212.1GFL. | | | |
| DR | GO; GO:0006091; P:energy pathways; IEA. | | | |
| DR | InterPro; IPR009017; GFP_like. | | | |
| DR | InterPro; IPR011584; GFP-related. | | | |
| DR | Pfam; PF01353; GFP. 1. | | | |
| DR | PRINTS; PR01229; GFP UOESCENT. | | | |
| DR | Prodom; PD013756; Green_fl_protein; 1. | | | |
| SO | SEQUENCE 238 AA; 26514 MW; P84840F1P9064018 CRC64; | | | |

| | | | | |
|---------------------------|--------|-------------------|------|------------------|
| Query Match | 100.0% | Score 1260 | DB 2 | Length 238 |
| Best Local Similarity | 98.3% | Pred. No. 3.1e-89 | | |
| Matches 234; Conservative | 4 | Mismatches | 0 | Indels 0; Gaps 0 |

| | | | |
|----|-----|---|-----|
| Qy | 1 | MSKGEELFTGVPVLPVLELDGDNVGNHSPVSGEEDGATGKLTIKICITGKGKPVMPPTL | 60 |
| Db | 1 | MSKGEELFTGVPVLPVLELDGDNVGNHSPVSGEEDGATGKLTIKICITGKGKPVMPPTL | 60 |
| Qy | 61 | VTTKXVYVQCFSRKPRDHMKHDFPKSAMPEGVQERTIEFKDGNKTRAEVFEEDTLV | 120 |
| Db | 61 | VTTKXVYVQCFSRKPRDHMKHDFPKSAMPEGVQERTIEFKDGNKTRAEVFEEDTLV | 120 |
| Qy | 121 | NRIELKSIDKEDGNILGHKLEYNNSHHVYIMADQKNGIKVNFRI RHINIEDGAVQLAD | 180 |
| Db | 121 | NRIELKSIDKEDGNILGHKLEYNNSHHVYIMADQKNGIKVNFRI RHINIEDGAVQLAD | 180 |
| Qy | 181 | HYQGNFTIGGGPVLLPDNNHLSLQSLSLSDPNKRDHMLLTPVTAAGITGHGDELYK | 238 |
| Db | 181 | HYQGNFTIGGGPVLLPDNNHLSLQSLSLSDPNKRDHMLLTPVTAAGITGHGDELYK | 238 |

| RESULT 2 | |
|-----------|---------------------------------|
| GFP_AEQVI | |
| ID | GFP_AEQVI STANDARD; PRT; 238 AA |
| AC | P42212; Q17104; Q27903; |

DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Green fluorescent protein.
GN Name=GFP.
OS Aequorea victoria (Jellyfish).
OC Eukaryota; Metazoa; Chnidaria; Hydrozoa; Hydrozoa; Leptomedusae;
OC Aequoreidae; Aequorea.
NCBI_Taxid=6100;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE=92175527; Pubmed=1347277; DOI=10.1016/0378-1119(92)90691-H;
RA Praeger D.C., Eckenrode V.K., Ward W.W., Prendergast F.G.,
RA Cormier M.J.;
RT "Primary structure of the Aequorea victoria green-fluorescent
RT protein.";
RL Gene 111:229-233 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94185810; Pubmed=8137953; DOI=10.1016/0014-5793(94)80472-9;
RA Inouye S., Tsuji F.I.;
RT "Aequorea green fluorescent protein. Expression of the gene and
RT fluorescence characteristics of the recombinant protein.";
RL FEBS Lett. 341:277-280 (1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97299832; Pubmed=9154981; DOI=10.1023/A:1005740823703;
RA Rowland J.J.A., Mendes O., Wolbert E.J.H., de Boer A.D.;
RT "Enhanced expression in tobacco of the gene encoding green fluorescent
RT protein by modification of its codon usage.";
RL Plant Mol. Biol. 33:989-999 (1997).
RN [4]
RP CHROMOPHORE.
RX MEDLINE=93192221; Pubmed=8448132;
RA Cody C.W., Praeger D.C., Westler W.M., Prendergast F.G., Ward W.W.;
RT "Chemical structure of the hexapeptide chromophore of the Aequorea
RT green-fluorescent protein.";
RL Biochemistry 32:1212-1218 (1993).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=96355665; Pubmed=8703075;
RA Ormoe W., Cubitt A.B., Kallio K., Gross L.A., Tsien R.Y.,
RA Remington S.J.;
RT "Crystal structure of the Aequorea victoria green fluorescent
RT protein.";
RL Science 273:1392-1395 (1996).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=98294543; Pubmed=9631087;
RA Yang F., Moss L.G., Phillips G.N., Jr.;
RT "The molecular structure of green fluorescent protein.";
RL Nat. Biotechnol. 14:1246-1251 (1996).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF MUTANT WITH YELLOW EMISSION.
RX MEDLINE=98455509; Pubmed=9782051; DOI=10.1016/S0969-2126(98)00127-0;
RA Wachter R.M., Bisliger M.A., Kallio K., Hanson G.T., Remington S.J.;
RT "Structural basis of spectral shifts in the yellow-emission variants
RT of green fluorescent protein.";
RL Structure 6:1267-1277 (1998).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=9928303; Pubmed=10220315; DOI=10.1021/bi9902182;
RA Bisliger M.A., Wachter R.M., Hanson G.T., Kallio K., Remington S.J.;
RT "Structural and spectral response of green fluorescent protein
RT variants to changes in pH.";
RL Biochemistry 38:5296-5301 (1999).
CC -1- FUNCTION: Energy-transfer acceptor. Its role is to transduce the
CC blue chemiluminescence of the protein aequorin into green
CC fluorescent light by energy transfer. Fluoresces in vivo upon
CC receiving energy from the Ca(2+)-activated photoprotein aequorin.
CC Absorbs light maximally at 395 nm and exhibits a smaller
CC absorbance peak at 470 nm. The fluorescence emission spectrum
CC peaks at 509 nm with a shoulder at 540 nm.

CC -1- SUBUNIT: Monomer.
CC -1- TISSUE SPECIFICITY: Photocytes.
CC -1- PTM: Contains a covalently attached chromophore, which is composed
CC of modified amino acid residues. The chromophore is formed upon
CC cyclization of the residues Ser-dehydrotyr-Gly.
CC -1- BIOTECHNOLOGY: Has become a useful and ubiquitous tool for making
CC chimeric proteins of GFP linked to other proteins where it
CC functions as a fluorescent protein tag. GFP tolerates N- and C-
CC terminal fusion to a broad variety of proteins. It has been
CC expressed in bacteria, yeast, slime mold, plants, Drosophila,
CC zebrafish, and in mammalian cells. As a noninvasive fluorescent
CC marker in living cells, it allows for a wide range of applications
CC where it may function as a cell lineage tracer, reporter of gene
CC expression, or as a measure of protein-protein interactions.
CC -1- DATABASE: NAME=Protein Spotlight; NOTE=Issue 11 of June 2001;
CC WWW=<http://www.expasy.org/spotlight/articles/aplt1011.html>.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M62654; AAA27722.1; -
DR EMBL; M62653; AAA27721.1; -
DR EMBL; L29345; AAA58246.1; -
DR EMBL; X96418; CAA65278.1; -
DR PIR; J050692; J01514.
DR PDB; 1B9C; X-ray; A/B/C/D=1-238.
DR PDB; 1BFP; X-ray; @=1-238.
DR PDB; 1CAF; X-ray; A=1-238.
DR PDB; 1CV7; X-ray; A=1-228.
DR PDB; 1EMA; X-ray; @=-1-238.
DR PDB; 1EMB; X-ray; @=1-238.
DR PDB; 1EMC; X-ray; A/B/C/D=1-238.
DR PDB; 1EME; X-ray; @=1-238.
DR PDB; 1EMF; X-ray; @=1-238.
DR PDB; 1EMG; X-ray; A=1-238.
DR PDB; 1EMK; X-ray; @=1-238.
DR PDB; 1EMN; X-ray; @=1-238.
DR PDB; 1EP9; X-ray; A=1-238.
DR PDB; 1F09; X-ray; A=1-238.
DR PDB; 1F0B; X-ray; A=1-238.
DR PDB; 1GFL; X-ray; A/B=1-238.
DR PDB; 1H8R; X-ray; A/B/C=1-238.
DR PDB; 1HCU; X-ray; A/B/C/D=1-238.
DR PDB; 1HUY; X-ray; A=1-238.
DR PDB; 1JBY; X-ray; A=1-238.
DR PDB; 1JH2; X-ray; A=1-238.
DR PDB; 1JCO; X-ray; A/B/C=1-238.
DR PDB; 1JCL; X-ray; A/B=1-237.
DR PDB; 1KP5; X-ray; A/B=1-237.
DR PDB; 1KYP; X-ray; A=1-238.
DR PDB; 1KXR; X-ray; A=1-238.
DR PDB; 1KYS; X-ray; A=1-238.
DR PDB; 1MTW; X-ray; A=1-238.
DR PDB; 1Q4A; X-ray; A=1-238.
DR PDB; 1Q4B; X-ray; A=1-238.
DR PDB; 1Q4D; X-ray; A=1-238.
DR PDB; 1Q4E; X-ray; A=1-238.
DR PDB; 1Q73; X-ray; A=1-229.
DR PDB; 1QXT; X-ray; A=1-229.
DR PDB; 1QYF; X-ray; A=1-229.
DR PDB; 1QYQ; X-ray; A=1-238.
DR PDB; 1QYR; X-ray; A=1-238.
DR PDB; 1YFP; X-ray; A/B=3-228.
DR PDB; 2EMD; X-ray; @=1-238.
DR PDB; 2EMN; X-ray; @=1-238.
DR PDB; 2EMO; X-ray; @=1-238.

| DR | InterPro: IPR009017; GFP-like. | |
|--|---|----------------------------|
| DR <td>InterPro: IPR000786; Green_fl_protein.</td> <td></td> | InterPro: IPR000786; Green_fl_protein. | |
| DR <td>Pfam: PF01353; GFP, 1.</td> <td></td> | Pfam: PF01353; GFP, 1. | |
| DR <td>PRINTS: PR01229; GFPUNRESSENT.</td> <td></td> | PRINTS: PR01229; GFPUNRESSENT. | |
| DR <td>ProDom: PD013756; Green_fl_protein; 1.</td> <td></td> | ProDom: PD013756; Green_fl_protein; 1. | |
| KW <td>3D-structure; Direct protein sequencing; Luminescence.</td> <td></td> | 3D-structure; Direct protein sequencing; Luminescence. | |
| FT <td>CROSSLINK 65 67</td> <td>5-indazoleone (Ser-Gly).</td> | CROSSLINK 65 67 | 5-indazoleone (Ser-Gly). |
| FT <td>MOD_RSS 66 66</td> <td>(Z)-2,3-didehydrotyrosine.</td> | MOD_RSS 66 66 | (Z)-2,3-didehydrotyrosine. |
| FT <td>VARIANT 100 100</td> <td>F -> Y.</td> | VARIANT 100 100 | F -> Y. |
| FT <td>VARIANT 108 108</td> <td>T -> S.</td> | VARIANT 108 108 | T -> S. |
| FT <td>VARIANT 141 141</td> <td>L -> M.</td> | VARIANT 141 141 | L -> M. |
| FT <td>VARIANT 219 219</td> <td>V -> I.</td> | VARIANT 219 219 | V -> I. |
| FT <td>CONFLICT 2 2</td> <td>S -> G (in Ref. 3).</td> | CONFLICT 2 2 | S -> G (in Ref. 3). |
| FT <td>CONFLICT 25 25</td> <td>H -> Q (in Ref. 2).</td> | CONFLICT 25 25 | H -> Q (in Ref. 2). |
| FT <td>CONFLICT 80 80</td> <td>Q -> R (in Ref. 3).</td> | CONFLICT 80 80 | Q -> R (in Ref. 3). |
| FT <td>CONFLICT 157 157</td> <td>Q -> P (in Ref. 2).</td> | CONFLICT 157 157 | Q -> P (in Ref. 2). |
| FT <td>CONFLICT 172 172</td> <td>E -> K (in Ref. 2).</td> | CONFLICT 172 172 | E -> K (in Ref. 2). |
| FT <td>TURN 3 3</td> <td></td> | TURN 3 3 | |
| FT <td>HELIX 4 8</td> <td></td> | HELIX 4 8 | |
| FT <td>STRAND 11 22</td> <td></td> | STRAND 11 22 | |
| FT <td>TURN 23 24</td> <td></td> | TURN 23 24 | |
| FT <td>STRAND 25 36</td> <td></td> | STRAND 25 36 | |
| FT <td>HELIX 37 39</td> <td></td> | HELIX 37 39 | |
| FT <td>TURN 40 40</td> <td></td> | TURN 40 40 | |
| FT <td>STRAND 41 48</td> <td></td> | STRAND 41 48 | |
| FT <td>TURN 49 50</td> <td></td> | TURN 49 50 | |
| FT <td>HELIX 57 60</td> <td></td> | HELIX 57 60 | |
| FT <td>TURN 61 63</td> <td></td> | TURN 61 63 | |
| FT <td>HELIX 69 71</td> <td></td> | HELIX 69 71 | |
| FT <td>STRAND 73 73</td> <td></td> | STRAND 73 73 | |
| FT <td>HELIX 76 81</td> <td></td> | HELIX 76 81 | |
| FT <td>HELIX 83 86</td> <td></td> | HELIX 83 86 | |
| FT <td>TURN 87 90</td> <td></td> | TURN 87 90 | |
| FT <td>STRAND 92 100</td> <td></td> | STRAND 92 100 | |
| FT <td>TURN 101 102</td> <td></td> | TURN 101 102 | |
| FT <td>STRAND 105 115</td> <td></td> | STRAND 105 115 | |
| FT <td>TURN 116 117</td> <td></td> | TURN 116 117 | |
| FT <td>STRAND 118 128</td> <td></td> | STRAND 118 128 | |
| FT <td>TURN 132 133</td> <td></td> | TURN 132 133 | |
| FT <td>TURN 135 139</td> <td></td> | TURN 135 139 | |
| FT <td>STRAND 141 141</td> <td></td> | STRAND 141 141 | |
| FT <td>STRAND 148 155</td> <td></td> | STRAND 148 155 | |
| FT <td>HELIX 156 158</td> <td></td> | HELIX 156 158 | |
| FT <td>TURN 159 159</td> <td></td> | TURN 159 159 | |
| FT <td>STRAND 160 171</td> <td></td> | STRAND 160 171 | |
| FT <td>TURN 172 173</td> <td></td> | TURN 172 173 | |
| FT <td>STRAND 176 187</td> <td></td> | STRAND 176 187 | |
| FT <td>STRAND 199 208</td> <td></td> | STRAND 199 208 | |
| FT <td>TURN 211 212</td> <td></td> | TURN 211 212 | |
| FT <td>STRAND 217 227</td> <td></td> | STRAND 217 227 | |
| SO <td>SEQUENCE 238 AA; 26886 MW; EA5A6F21FBA6E05 CRC64;</td> <td></td> | SEQUENCE 238 AA; 26886 MW; EA5A6F21FBA6E05 CRC64; | |
| Query Match | 99.7%; Score 1256; DB 1; Length 238; | |
| Best Local Similarity | 97.9%; Pred. No. 6,3e-89; | |
| Matches 233; Conservative | 5; Mismatches 0; Indels 0; Gaps 0; | |
| QY | 1 MSKGEELTGTVVPIIVELDGDVNGHKFSVSGEGDATYGKLTLEICTTGKLPVMPPTL 60 | |
| DB | 1 MSKGEELTGTVVPIIVELDGDVNGHKFSVSGEGDATYGKLTLEICTTGKLPVMPPTL 60 | |
| QY | 61 VTTXXYGVQCFSRYPDHNKRHDFKSAPEGVQERTIFFKODGNYKTRAEVKFEEDTLV 120 | |
| DB | 61 VTTTFYGVQCFSRYPDHNKRHDFKSAPEGVQERTIFFKODGNYKTRAEVKFEEDTLV 120 | |
| QY | 121 NRLEIKGIDFKEDGNIILGHKLEYNNSHNVYIMADOKNGIKVNFKIRANIEDGVQLAD 180 | |
| DB | 121 NRLEIKGIDFKEDGNIILGHKLEYNNSHNVYIMADOKNGIKVNFKIRANIEDGVQLAD 180 | |
| QY | 161 HYQONTPIGDGQVLLPDDNHYLSTQSALSQDPNEKDDHWLILKFTYAAGITGHMDELYK 238 | |
| DB | 161 HYQONTPIGDGQVLLPDDNHYLSTQSALSQDPNEKDDHWLILKFTYAAGITGHMDELYK 238 | |

| ID | Q71RY9 | PRELIMINARY; | PRT; | 238 AA. |
|---------------------------|---|---|------|---------|
| AC | Q71RY9 | | | |
| DT | 05-JUL-2004 | (TREMBLrel. 27, Created) | | |
| DT | 05-JUL-2004 | (TREMBLrel. 27, Last sequence update) | | |
| DT | 05-JUL-2004 | (TREMBLrel. 27, Last annotation update) | | |
| DE | Green fluorescence protein. | | | |
| GN | Name=2896fp; | | | |
| OS | Azotobacter vinelandii. | | | |
| OC | Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; | | | |
| OC | Pseudomonadaceae; Azotobacter. | | | |
| OX | NCBI_TaxID=354; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=CCM289; | | | |
| RA | Koranyi P., Berenyi M., Burg K.; | | | |
| RL | Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases. | | | |
| RL | EMBL; AF344407; AAN86139.1; -. | | | |
| DR | HSSP; P42212; 1B9C. | | | |
| DR | GO; GO:0006091; P:energy pathways; IEA. | | | |
| DR | InterPro; IPR009017; GFP_like. | | | |
| DR | InterPro; IPR011584; GFP-related. | | | |
| DR | InterPro; IPR000786; Green_fl_protein. | | | |
| DR | Pfam; PF01353; GFP. 1. | | | |
| DR | PRINTS; PR01229; GFPUNRESCENT. | | | |
| DR | Prodom; PD013756; Green_fl_protein; 1. | | | |
| SQ | SEQUENCE 238 AA; 26886 MW; EA5A6F21FBB6E05 CRC64; | | | |
| Query Match | 99.7%; Score 1256; DB 2; Length 238; | | | |
| Best Local Similarity | 97.9%; Pred. No. 6.3e-89; | | | |
| Matches 233; Conservative | 5; Mismatches 0; Indels 0; Gaps 0; | | | |
| OY | 1 MSKEELFTGVVPLVLEIDGDVNGHKESVSGEGGDATYGLTLKFICTTGKLPVMPPTL 60 | | | |
| DB | 1 MSKEELFTGVVPLVLEIDGDVNGHKESVSGEGGDATYGLTLKFICTTGKLPVMPPTL 60 | | | |
| OY | 61 VTTXXYGVQCSRRYPDHMKRHDFFKSAPEGYVOERTIFFKDDGNYKTRAEVKEGDTLV 120 | | | |
| DB | 61 VTTFSYGVQCSRRYPDHMKRHDFFKSAPEGYVOERTIFFKDDGNYKTRAEVKEGDTLV 120 | | | |
| OY | 121 NRIELKGI DFEKDNITLGHKLEVYNVNSHNYIMDKONGIKVNPKIRHNI EDGVQLAD 180 | | | |
| DB | 121 NRIELKGI DFEKDNITLGHKLEVYNVNSHNYIMDKONGIKVNPKIRHNI EDGVQLAD 180 | | | |
| OY | 181 HYQONTPIGDSPVLLPDNHVLTOSTALS KDPNEKRDMVLLXFTYAGITGMDLYK 238 | | | |
| DB | 181 HYQONTPIGDSPVLLPDNHVLTOSTALS KDPNEKRDMVLLXFTYAGITGMDLYK 238 | | | |
| RESULT 4 | | | | |
| ID | O8GHE4 | PRELIMINARY; | PRT; | 238 AA. |
| AC | O8GHE4 | | | |
| DT | 01-MAR-2003 | (TREMBLrel. 23, Created) | | |
| DT | 01-MAR-2003 | (TREMBLrel. 23, Last sequence update) | | |
| DT | 01-OCT-2003 | (TREMBLrel. 25, Last annotation update) | | |
| DE | Green fluorescence protein. | | | |
| GN | Name=3756fp; | | | |
| OS | Azomonas agilis. | | | |
| OC | Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; | | | |
| OC | Pseudomonadaceae; Azomonas. | | | |
| OX | NCBI_TaxID=116849; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | Koranyi P., Berenyi M., Burg K.; | | | |
| RL | Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases. | | | |
| RL | EMBL; AF344405; AAN86137.1; -. | | | |
| DR | HSSP; P42212; 1B9P. | | | |
| DR | GO; GO:0006091; P:energy pathways; IEA. | | | |
| DR | InterPro; IPR009017; GFP_like. | | | |
| DR | InterPro; IPR011584; GFP-related. | | | |
| DR | InterPro; IPR000786; Green_fl_protein. | | | |

DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFPLORESCENT.
DR Prodom; PD013756; Green fl protein; 1.
SQ SEQUENCE 238 AA; 26902 MW; 15FE9B9C5B4F6B89 CRC64;

Query Match 99.6%; Score 1255; DB 2; Length 238;
Best Local Similarity 97.9%; Pred. No. 7.5e-89;
Matches 233; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGSDATYGLTKLTKICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGSDATYGLTKLTKICTTGKLPVWPPTL 60
QY VTTXXYGVCCFSRYPDHMKRHDFFKSAPEGVQERTIFFKDGNYKTRAIVKFEEDTLV 120
DB 61 VTTFSYGVCCFSRYPDHMKRHDFFKSAPEGVQERTIFFKDGNYKTRAIVKFEEDTLV 120
QY 121 NRTELKIDFKEDGNILGHKLEYNNSHNYIMADKQKGIKYNFKIRHNIEDGVQLAD 180
DB 121 NRTELKIDFKEDGNILGHKLEYNNSHNYIMADKQKGIKYNFKIRHNIEDGVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLFTVTAAGITHGMDLYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLFTVTAAGITHGMDLYK 238

RESULT 5

ID 08GHR3 PRELIMINARY; PRT; 238 AA.
AC 08GHR3;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Green fluorescence protein.
GN Name=85Gfp;
OS Azotobacter vinelandii;
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_Taxid=354;
RN [1]
RP SEQUENCE FROM N.A.
RA Koranyi P., Berenyi M., Burg K.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF324406; AAN86138.1; -.
DR HSSP; P42212; 1GFL.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP_like.
DR InterPro; IPR011584; GFP_related.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFPLORESCENT.
DR Prodom; PD013756; Green fl protein; 1.
SQ SEQUENCE 238 AA; 26887 MW; E0B1616BD2AF6188 CRC64;

Query Match 99.5%; Score 1254; DB 2; Length 238;
Best Local Similarity 97.9%; Pred. No. 9e-89;
Matches 233; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGSDATYGLTKLTKICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGSDATYGLTKLTKICTTGKLPVWPPTL 60
QY VTTXXYGVCCFSRYPDHMKRHDFFKSAPEGVQERTIFFKDGNYKTRAIVKFEEDTLV 120
DB 61 VTTFSYGVCCFSRYPDHMKRHDFFKSAPEGVQERTIFFKDGNYKTRAIVKFEEDTLV 120
QY 121 NRTELKIDFKEDGNILGHKLEYNNSHNYIMADKQKGIKYNFKIRHNIEDGVQLAD 180
DB 121 NRTELKIDFKEDGNILGHKLEYNNSHNYIMADKQKGIKYNFKIRHNIEDGVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLFTVTAAGITHGMDLYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLFTVTAAGITHGMDLYK 238

RESULT 6

ID 093125 PRELIMINARY; PRT; 238 AA.
AC 093125;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Green fluorescent protein mutant 3.
GN Name=GFP;
OS Aequorea victoria (Uelilyfish).
OC Eukaryota; Metazoa; Chordata; Hydrozoa; Hydrozoa; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_Taxid=6100;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96305137; PubMed=8707053; DOI=10.1016/0378-1119(95)00685-0;
RA Cornack B.P., Valdivia R.H., Falkow S.;
RT "FACS-optimized mutants of the green fluorescent protein (GFP).";
RL Gene 173:33-38(1996).

QY 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGSDATYGLTKLTKICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGSDATYGLTKLTKICTTGKLPVWPPTL 60
QY VTTXXYGVCCFSRYPDHMKRHDFFKSAPEGVQERTIFFKDGNYKTRAIVKFEEDTLV 120
DB 61 VTTFSYGVCCFSRYPDHMKRHDFFKSAPEGVQERTIFFKDGNYKTRAIVKFEEDTLV 120
QY 121 NRTELKIDFKEDGNILGHKLEYNNSHNYIMADKQKGIKYNFKIRHNIEDGVQLAD 180
DB 121 NRTELKIDFKEDGNILGHKLEYNNSHNYIMADKQKGIKYNFKIRHNIEDGVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLFTVTAAGITHGMDLYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLFTVTAAGITHGMDLYK 238

Query Match 99.4%; Score 1253; DB 2; Length 238;
Best Local Similarity 97.5%; Pred. No. 1.1e-88;
Matches 232; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGSDATYGLTKLTKICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLELDGVNKHKFSVSGEGSDATYGLTKLTKICTTGKLPVWPPTL 60
QY VTTXXYGVCCFSRYPDHMKRHDFFKSAPEGVQERTIFFKDGNYKTRAIVKFEEDTLV 120
DB 61 VTTFSYGVCCFSRYPDHMKRHDFFKSAPEGVQERTIFFKDGNYKTRAIVKFEEDTLV 120
QY 121 NRTELKIDFKEDGNILGHKLEYNNSHNYIMADKQKGIKYNFKIRHNIEDGVQLAD 180
DB 121 NRTELKIDFKEDGNILGHKLEYNNSHNYIMADKQKGIKYNFKIRHNIEDGVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLFTVTAAGITHGMDLYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLFTVTAAGITHGMDLYK 238

RESULT 7

ID 017105 PRELIMINARY; PRT; 238 AA.
AC 017105;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Green fluorescent protein (Fragment).
GN Name=GFP;
OS Aequorea victoria (Uelilyfish).
OC Eukaryota; Metazoa; Chordata; Hydrozoa; Hydrozoa; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_Taxid=6100;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Wackine J.N., Campbell A.K.;
 RL Submitted (Jan-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X83959; CAA58789.1; -.
 DR PIR; J50692; J01514.
 DR HSSP; P42212; 1GFL.
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR InterPro; IPR009017; GFP like.
 DR InterPro; IPR011584; GFP_related.
 DR InterPro; IPR000786; Green_fl_protein.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFP.1.
 DR FT NON_TER
 FT 238
 SQ SEQUENCE 238 AA; 26950 MW; 26E2BE450E748B44 CRC64;

Query Match 96.3%; Score 1214; DB 2; Length 238;
 Best Local Similarity 93.7%; Pred. No. 1.1e-85;
 Matches 223; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGELFTGVVPIIVELDGDVNGHKFSVSGEGEDATYTKLFTCTTGKLPVPWPTL 60
 DB 1 MSKGELFTGVVPIIVELDGDVNGKFSVSGEGEDATYTKLFTCTTGKLPVPWPTL 60
 QY 61 VTTXXYGVQCFSRYPDHHKRDHDFKSAPEGVQERTIFFKDDGNYKTRAEVKFEEDTLV 120
 DB 61 VTTFSYGVQCFSRYPDHHKRDHDFKSAPEGVQERTIFFKDDGNYKTRAEVKFEEDTLV 120
 QY 121 NRIELKGIQKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGVQLAD 180
 DB 121 NRIELKGIQKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGVQLAD 180
 QY 121 NRIELKGIQKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGVQLAD 180
 DB 121 NRIELKGIQKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGVQLAD 180
 QY 181 HYQONTPIGDGVPVLLPDNHYLSTQSALSKDPNEKRDHVLKFTVTAAGITTHGMDELTX 238
 DB 181 HYQONTPIGDGVPVLLPDNHYLSTQSALSKDPNEKRDHVLKFTVTAAGITTHGMDELTX 238

RESULT 8

ID 017106 PRELIMINARY; PRT; 238 AA.
 AC 017106;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Green fluorescent protein (fragment).
 GN Name=GFP;
 OS Aequorea victoria (Jellyfish).
 OC Eukaryota; Metazoa; Chordata; Hydrozoa; Hydrozoa; Leptomedusae;
 CC Aequoreidae; Aequorea.
 CX NCBI_TaxID=6100;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wackine J.N., Campbell A.K.;
 RL Submitted (Jan-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X83960; CAA58790.1; -.
 DR PIR; J50693; J01514.
 DR HSSP; P42212; 1BFP.
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR InterPro; IPR009017; GFP like.
 DR InterPro; IPR011584; GFP_related.
 DR InterPro; IPR000786; Green_fl_protein.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFP.1.
 DR FT NON_TER
 FT 238
 SQ SEQUENCE 238 AA; 26867 MW; BD4648262D8EABD4 CRC64;

Query Match 95.2%; Score 1199; DB 2; Length 238;
 Best Local Similarity 92.9%; Pred. No. 1.6e-84;
 Matches 221; Conservative 11; Mismatches 6; Indels 0; Gaps 0;

QY 1 MSKGELFTGVVPIIVELDGDVNGHKFSVSGEGEDATYTKLFTCTTGKLPVPWPTL 60
 DB 1 MSKGELFTGVVPIIVELDGDVNGKFSVSGEGEDATYTKLFTCTTGKLPVPWPTL 60

QY 61 VTTXXYGVQCFSRYPDHHKRDHDFKSAPEGVQERTIFFKDDGNYKTRAEVKFEEDTLV 120
 DB 61 VTTFSYGVQCFSRYPDHHKRDHDFKSAPEGVQERTIFFKDDGNYKTRAEVKFEEDTLV 120
 QY 121 NRIELKGIQKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGVQLAD 180
 DB 121 NRIELKGIQKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGVQLAD 180
 QY 121 NRIELKGIQKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGVQLAD 180
 DB 121 NRIELKGIQKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGVQLAD 180
 QY 181 HYQONTPIGDGVPVLLPDNHYLSTQSALSKDPNEKRDHVLKFTVTAAGITTHGMDELTX 238
 DB 181 HYQONTPIGDGVPVLLPDNHYLSTQSALSKDPNEKRDHVLKFTVTAAGITTHGMDELTX 238

RESULT 9

ID 06YG20 PRELIMINARY; PRT; 238 AA.
 AC 06YG20;
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Green fluorescent protein.
 OS Aequorea coerulescens (belt jellyfish).
 OC Eukaryota; Metazoa; Chordata; Hydrozoa; Hydrozoa; Leptomedusae;
 CC Aequoreidae; Aequorea.
 CX NCBI_TaxID=210840;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=22726112; PubMed=12693991; DOI=10.1042/BJ20021966;
 RA Gurekaya N.G., Fradkov A.F., Pounkova N.I., Starovtsev D.B.,
 RA Bulina M.R., Yanshevich Y.G., Labas Y.A., Lukyanov S., Lukyanov K.A.;
 RT "A colourless green fluorescent protein homologue from the non-fluorescent hydromedusa Aequorea coerulescens and its fluorescent mutants";
 RL Biochem. J. 373:403-408(2003).
 DR EMBL; AY151052; AAN41637.1; -.
 DR HSSP; P42212; 1B9C.
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR InterPro; IPR009017; GFP like.
 DR InterPro; IPR011584; GFP_related.
 DR InterPro; IPR000786; Green_fl_protein.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFP.1.
 DR PRODOM; PD013756; Green_fl_protein; 1.
 SQ SEQUENCE 238 AA; 26856 MW; DE72EDB87ED99FE CRC64;

Query Match 94.1%; Score 1186; DB 2; Length 238;
 Best Local Similarity 90.8%; Pred. No. 1.6e-83;
 Matches 216; Conservative 13; Mismatches 9; Indels 0; Gaps 0;

QY 1 MSKGELFTGVVPIIVELDGDVNGHKFSVSGEGEDATYTKLFTCTTGKLPVPWPTL 60
 DB 1 MSKGELFTGVVPIIVELDGDVNGHKFSVSGEGEDATYTKLFTCTTGKLPVPWPTL 60
 QY 61 VTTXXYGVQCFSRYPDHHKRDHDFKSAPEGVQERTIFFKDDGNYKTRAEVKFEEDTLV 120
 DB 61 VTTFSYGVQCFSRYPDHHKRDHDFKSAPEGVQERTIFFKDDGNYKTRAEVKFEEDTLV 120
 QY 121 NRIELKGIQKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGVQLAD 180
 DB 121 NRIELKGIQKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGVQLAD 180
 QY 121 NRIELKGIQKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGVQLAD 180
 DB 121 NRIELKGIQKEDGNILGHKLEYNNSHNVYIMADKONGIKVNFKIRHNIEDGVQLAD 180
 QY 181 HYQONTPIGDGVPVLLPDNHYLSTQSALSKDPNEKRDHVLKFTVTAAGITTHGMDELTX 238
 DB 181 HYQONTPIGDGVPVLLPDNHYLSTQSALSKDPNEKRDHVLKFTVTAAGITTHGMDELTX 238

RESULT 10

ID 08WP95 PRELIMINARY; PRT; 238 AA.
 AC 08WP95;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DT 05-JUL-2004 (T-EMBLrel. 27, last annotation update)
 DE Green fluorescent protein.
 GN Name=GFpam;
 OS Aequorea macrodactyla.
 OC Eukaryota; Metazoa; Chordata; Hydrozoa; Hydroida; Leptomedusae;
 OC Aequoreidae; Aequorea.
 RN NCBL_TaxID=147615;
 RP SEQUENCE FROM N.A.
 RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Qin Y.X., Pang S.Q.,
 RA Li S.J., Xia N.S.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY013824; AK02062.1; -
 DR EMBL; AY013821; AK02059.1; -
 DR HSSP; P42212; 1BFP
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFP/DRESCENT.
 SQ SEQUENCE 238 AA; 27049 MW; 8185D0E5E529012B CRC64;

Query Match 87.1%; Score 1097; DB 2; Length 238;
 Best Local Similarity 82.4%; Pred. No. 1.1e-76;
 Matches 196; Conservative 23; Mismatches 19; Indels 0; Gaps 0;

QY 1 MSKGBELFTGVPIVLELDGVDVNGHKFSVSGEGSDATYGTTLKFTCTTGKLPVWPPTL 60
 DB 1 MSKGBELFTGVPIVLELDGVDVNGHKFSVSGEGSDATYGTTLKFTCTTGKLPVWPPTL 60
 QY 61 VTTXXYGVOCFSRYPDHMKRHDFPKSAMPEGYVOERTIFFKDGNYKTRAEVYFEGDTLY 120
 DB 61 VTTFSYGIQCFARYPEHMKRNDFFKSAMPEGYIOERTIFFODDGKXKTRGEVYFEGDTLY 120
 QY 121 NRTELKIDFKEDGNILGHKLEYNVNSHNVYIMADKQNGIKVNFKIRHNIEDGVQVLAD 180
 DB 121 NRTELKMDFKEDGNILGHKLEYNFNSHNVYIMPDKANNGLKVNFKIRHNIEGGGVQVLAD 180
 QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLXFTVTAAGITTHGMDELYK 238
 DB 181 HYQTNVPLDGPVLLPINHYLSTQTALSKDRNETRDHMLVLEFFSACGTHGMDELYK 238

RESULT 11

Q8WTC6 PRELIMINARY; PRT; 238 AA.
 ID Q8WTC6
 AC Q8WTC6;
 DT 01-MAR-2002 (T-EMBLrel. 20, Created)
 DT 01-MAR-2002 (T-EMBLrel. 20, last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, last annotation update)
 DE Green fluorescent protein.
 GN Name=GFP;
 OS Aequorea macrodactyla.
 OC Eukaryota; Metazoa; Chordata; Hydrozoa; Hydroida; Leptomedusae;
 OC Aequoreidae; Aequorea.
 RN NCBL_TaxID=147615;
 RP SEQUENCE FROM N.A.
 RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
 RA Li S.J., Xia N.S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF435431; AL33916.1; -
 DR HSSP; P42212; 1KYP.
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFP/DRESCENT.
 SQ SEQUENCE 238 AA; 27015 MW; 6B8FD75EB8926903 CRC64;

Query Match 87.1%; Score 1097; DB 2; Length 238;
 Best Local Similarity 82.4%; Pred. No. 1.1e-76;
 Matches 196; Conservative 23; Mismatches 19; Indels 0; Gaps 0;

QY 1 MSKGBELFTGVPIVLELDGVDVNGHKFSVSGEGSDATYGTTLKFTCTTGKLPVWPPTL 60
 DB 1 MSKGBELFTGVPIVLELDGVDVNGHKFSVSGEGSDATYGTTLKFTCTTGKLPVWPPTL 60

QY 61 VTTXXYGVOCFSRYPDHMKRHDFPKSAMPEGYVOERTIFFKDGNYKTRAEVYFEGDTLY 120
 DB 61 VTTLSYGIQCFARYPEHMKRNDFFKSAMPEGYIOERTIFFODDGKXKTRGEVYFEGDTLY 120
 QY 121 NRTELKIDFKEDGNILGHKLEYNVNSHNVYIMADKQNGIKVNFKIRHNIEDGVQVLAD 180
 DB 121 NRTELKMDFKEDGNILGHKLEYNFNSHNVYIMPDKANNGLKVNFKIRHNIEGGGVQVLAD 180
 QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLXFTVTAAGITTHGMDELYK 238
 DB 181 HYQTNVPLDGPVLLPINHYLSTQTALSKDRNETRDHMLVLEFFSACGTHGMDELYK 238

RESULT 12

Q8WTC4 PRELIMINARY; PRT; 238 AA.
 ID Q8WTC4
 AC Q8WTC4;
 DT 01-MAR-2002 (T-EMBLrel. 20, Created)
 DT 01-MAR-2002 (T-EMBLrel. 20, last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, last annotation update)
 DE Green fluorescent protein.
 GN Name=GFP;
 OS Aequorea macrodactyla.
 OC Eukaryota; Metazoa; Chordata; Hydrozoa; Hydroida; Leptomedusae;
 OC Aequoreidae; Aequorea.
 RN NCBL_TaxID=147615;
 RP SEQUENCE FROM N.A.
 RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
 RA Li S.J., Xia N.S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF435433; AL33918.1; -
 DR HSSP; P42212; 1KYP.
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFP/DRESCENT.
 SQ SEQUENCE 238 AA; 26956 MW; 75521EAF0CEBA73A CRC64;

Query Match 86.7%; Score 1093; DB 2; Length 238;
 Best Local Similarity 83.2%; Pred. No. 2.3e-76;
 Matches 198; Conservative 20; Mismatches 20; Indels 0; Gaps 0;

QY 1 MSKGBELFTGVPIVLELDGVDVNGHKFSVSGEGSDATYGTTLKFTCTTGKLPVWPPTL 60
 DB 1 MSKGBELFTGVPIVLELDGVDVNGHKFSVSGEGSDATYGTTLKFTCTTGKLPVWPPTL 60
 QY 61 VTTXXYGVOCFSRYPDHMKRHDFPKSAMPEGYVOERTIFFKDGNYKTRAEVYFEGDTLY 120
 DB 61 VTTLSYGIQCFARYPEHMKRNDFFKSAMPEGYIOERTIFFODDGKXKTRGEVYFEGDTLY 120
 QY 121 NRTELKIDFKEDGNILGHKLEYNVNSHNVYIMADKQNGIKVNFKIRHNIEDGVQVLAD 180
 DB 121 NRTELKMDFKEDGNILGHKLEYNFNSHNVYIMPDKANNGLKVNFKIRHNIEGGGVQVLAD 180
 QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLXFTVTAAGITTHGMDELYK 238
 DB 181 HYQTNVPLDGPVLLPINHYLSTQTALSKDRNETRDHMLVLEFFSACGTHGMDELYK 238

RESULT 13

Q8WTD0 PRELIMINARY; PRT; 238 AA.
 ID Q8WTD0
 AC Q8WTD0;
 DT 01-MAR-2002 (T-EMBLrel. 20, Created)
 DT 01-MAR-2002 (T-EMBLrel. 20, last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, last annotation update)
 DE Green fluorescent protein.
 GN Name=GFP;
 OS Aequorea macrodactyla.
 OC Eukaryota; Metazoa; Chordata; Hydrozoa; Hydroida; Leptomedusae;
 OC Aequoreidae; Aequorea.
 RN NCBL_TaxID=147615;

[1]
 RN SEQUENCE FROM N.A.
 RP Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.O., Chen M.,
 RA Li S.J., Xia N.S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF435427; AAL3912.1; -.
 DR HSSP; P42212; IKYP.
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFP; 1.
 DR PRODOM; PD013756; Green_Fl_protein; 1.
 SQ SEQUENCE 238 AA; 26997 MW; 5F80A192173C884D CRC64;

Query Match 86.6%; Score 1091; DB 2; Length 238;
 Best Local Similarity 81.9%; Pred. No. 33e-76;
 Matches 195; Conservative 23; Mismatches 20; Indels 0; Gaps 0;

QY 1 MSKGELFTGVVPIVELDGVNGHKFSVSGEGDATYGLTLKFICTTGKLPVWPPTL 60
 DB 1 MSKGELFTGIVPVLIELDGDVGHKFSVRGEGDADYGLKIKFICTTGKLPVWPPTL 60
 QY 61 VTTXXYGVQCFSRYPDHMKRHDFFKSAMEGYVOERTIFFKDDGNYKTRAEVKFEEDTLV 120
 DB 61 VTTLCYGIQCFARYBEHMKMDFFKSAMEGYIOERTIFFODGKYKTRGEVYFEGDTLV 120
 QY 121 NRIELKGIDFKEDGNILGHKLEYNNSHVYIMADKQKGIIVNFKIRINIEDGVYQLAD 180
 DB 121 NRIELKGMDFKEDGNILGHKLEYNNSHVYIMPDKANGLVNFKIRINIEGGVQLAD 180
 QY 181 HYQONTPIGDGPVLLPDNHVYSTQSALSQDPNEKRDHVLTKFVTPAAGITTHGMDELYK 238
 DB 181 HYQTNVPLGDGPVLLPINHYLSQTALSKDRNETRDHNVLFEPFSA CGHTHGMDELYK 238

RESULT 14

Q8WTC8 PRELIMINARY; PRT; 238 AA.

ID 08WTC8
 AC 08WTC8;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Green fluorescent protein.
 GN Name=GFP;
 OS Aequorea macrodactyla.
 OC Eukaryota; Metazoa; Chordata; Hydrozoa; Hydrozoa; Leptomedusae;
 OC Aequoreidae; Aequorea.
 OX NCBI_TaxId=147615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.O., Chen M.,
 RA Li S.J., Xia N.S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF435429; AAL3914.1; -.
 DR HSSP; P42212; IKYP.
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFP; 1.
 SQ SEQUENCE 238 AA; 27047 MW; 5F80A18FA1E7C84D CRC64;

Query Match 86.5%; Score 1090; DB 2; Length 238;
 Best Local Similarity 81.9%; Pred. No. 4e-76;
 Matches 195; Conservative 23; Mismatches 20; Indels 0; Gaps 0;

QY 1 MSKGELFTGVVPIVELDGVNGHKFSVSGEGDATYGLTLKFICTTGKLPVWPPTL 60
 DB 1 MSKGELFTGIVPVLIELDGDVGHKFSVRGEGDADYGLKIKFICTTGKLPVWPPTL 60
 QY 61 VTTXXYGVQCFSRYPDHMKRHDFFKSAMEGYVOERTIFFKDDGNYKTRAEVKFEEDTLV 120
 DB 61 VTTLCYGIQCFARYBEHMKMDFFKSAMEGYIOERTIFFODGKYKTRGEVYFEGDTLV 120
 QY 121 NRIELKGIDFKEDGNILGHKLEYNNSHVYIMADKQKGIIVNFKIRINIEDGVYQLAD 180

DB 121 NRIELKGMDFKEDGNILGHKLEYNNSHVYIMPDKANGLVNFKIRINIEGGVQLAD 180
 QY 181 HYQONTPIGDGPVLLPDNHVYSTQSALSQDPNEKRDHVLTKFVTPAAGITTHGMDELYK 238
 DB 181 HYQTNVPLGDGPVLLPINHYLSQTALSKDRNETRDHNVLFEPFSA CGHTHGMDELYK 238

RESULT 15

Q8WTC9 PRELIMINARY; PRT; 238 AA.

ID 08WTC9
 AC 08WTC9;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Green fluorescent protein.
 GN Name=GFP;
 OS Aequorea macrodactyla.
 OC Eukaryota; Metazoa; Chordata; Hydrozoa; Hydrozoa; Leptomedusae;
 OC Aequoreidae; Aequorea.
 OX NCBI_TaxId=147615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.O., Chen M.,
 RA Li S.J., Xia N.S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF435428; AAL3913.1; -.
 DR HSSP; P42212; IKYP.
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFP; 1.
 SQ SEQUENCE 238 AA; 27031 MW; 5F80A19C19DC884D CRC64;

Query Match 86.5%; Score 1090; DB 2; Length 238;
 Best Local Similarity 81.9%; Pred. No. 4e-76;
 Matches 195; Conservative 23; Mismatches 20; Indels 0; Gaps 0;

QY 1 MSKGELFTGVVPIVELDGVNGHKFSVSGEGDATYGLTLKFICTTGKLPVWPPTL 60
 DB 1 MSKGELFTGIVPVLIELDGDVGHKFSVRGEGDADYGLKIKFICTTGKLPVWPPTL 60
 QY 61 VTTXXYGVQCFSRYPDHMKRHDFFKSAMEGYVOERTIFFKDDGNYKTRAEVKFEEDTLV 120
 DB 61 VTTLCYGIQCFARYBEHMKMDFFKSAMEGYIOERTIFFODGKYKTRGEVYFEGDTLV 120
 QY 121 NRIELKGIDFKEDGNILGHKLEYNNSHVYIMADKQKGIIVNFKIRINIEDGVYQLAD 180
 DB 121 NRIELKGMDFKEDGNILGHKLEYNNSHVYIMPDKANGLVNFKIRINIEGGVQLAD 180
 QY 181 HYQONTPIGDGPVLLPDNHVYSTQSALSQDPNEKRDHVLTKFVTPAAGITTHGMDELYK 238
 DB 181 HYQTNVPLGDGPVLLPINHYLSQTALSKDRNETRDHNVLFEPFSA CGHTHGMDELYK 238

Search completed: February 28, 2005, 14:29:31
 Job time : 176 secs

This Page Blank (uspio)